

STIC-Biotech/ChemLib

152404

mg

From: Kaushal, Sumesh  
Sent: Tuesday, May 03, 2005 5:45 PM  
To: STIC-Biotech/ChemLib  
Subject: 10657852: SEQ search

**10657852: SEQ search**

**Please search**

- SEQ ID NO:3
- SEQ ID NO:15

*S. Kaushal*

AU1636, REM2.B85  
Ph: 571-27-20769  
Mail Box: REM2.C70

RECEIVED  
MAY - 4 2005  
STIC-Biotech Division  
(STIC)

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: 5/6/05  
Date Completed: 5/11/05  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA#: 1 AA#: 1  
Interference: \_\_\_\_\_ SPDI: \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure#: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: out  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

*This Page Blank (uspto)*



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 152404

TO: Sumesh Kaushal  
Location: rem/2b85/2c70  
Art Unit: 1636  
Wednesday, May 11, 2005

Case Serial Number: 10/657852

From: Edward Hart  
Location: Biotech-Chem Library  
REM-1A55  
Phone: 571-272-2512

[edward.hart@uspto.gov](mailto:edward.hart@uspto.gov)

### Search Notes

Examiner Kaushal,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart



*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - protein search, using sw model

Run on: May 9, 2005, 20:12:52 ; Search time 65 Seconds  
(without alignments)  
2103.464 Million cell updates/sec

Title: US-10-657-852A-15

Perfect score: 1385

Sequence: 1 MPFYMAKCCMLLVFLGLFIQ.....NTVSGSNHIVSGSNKVVTDG 267

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID        | Description        |
|------------|-------|---------------|--------|--------------|--------------------|
| 1          | 530   | 38.3          | 118    | 2 Q9M3W4     | Q9m3w4 lolium pere |
| 2          | 262   | 18.9          | 1010   | 2 Q8H3W8     | Q8h3w8 oryza sativ |
| 3          | 258.5 | 18.7          | 1036   | 2 Q9FN37     | Q9fn37 arabidopsis |
| 4          | 221.5 | 16.0          | 1124   | 2 Q63UH1     | Q63uh1 burkholderi |
| 5          | 218   | 15.7          | 1012   | 2 Q62KN5     | Q62kn5 burkholderi |
| 6          | 212   | 15.3          | 333    | 2 Q9ZNX4     | Q9znx4 petunia hyb |
| 7          | 204   | 14.7          | 1021   | 1 PSKR DAUCA | Q8lbp4 daucus caro |
| 8          | 201   | 14.5          | 1065   | 2 Q66QA4     | Q66qa4 oryza sativ |
| 9          | 199.5 | 14.4          | 784    | 2 Q9C9H6     | Q9c9h6 arabidopsis |
| 10         | 197.5 | 14.3          | 751    | 2 Q940W7     | Q940w7 arabidopsis |
| 11         | 197.5 | 14.3          | 751    | 2 Q9FGQ5     | Q9fgq5 arabidopsis |
| 12         | 196   | 14.2          | 1102   | 2 Q69KC3     | Q69kc3 oryza sativ |
| 13         | 195.5 | 14.1          | 1023   | 2 Q84NG8     | Q84ng8 hordeum vul |
| 14         | 194   | 14.0          | 622    | 2 Q66CJ0     | Q66cj0 yersinia ps |
| 15         | 194   | 14.0          | 917    | 2 Q75GM9     | Q75gm9 oryza sativ |
| 16         | 194   | 14.0          | 1063   | 2 Q67IT7     | Q67it7 oryza sativ |
| 17         | 193.5 | 14.0          | 372    | 2 Q89J37     | Q89j37 bradyrhizob |
| 18         | 192.5 | 13.9          | 846    | 2 Q9C9H7     | Q9c9h7 arabidopsis |
| 19         | 192   | 13.9          | 727    | 2 Q67TW5     | Q67tw5 oryza sativ |
| 20         | 192   | 13.9          | 1060   | 2 Q6ZGM3     | Q6zgm3 oryza sativ |
| 21         | 190.5 | 13.8          | 943    | 2 Q9SRL7     | Q9srl7 arabidopsis |
| 22         | 190.5 | 13.8          | 1008   | 1 PSKR ARATH | Q9zvr7 arabidopsis |
| 23         | 190   | 13.7          | 622    | 2 Q8ZGB5     | Q8zgb5 yersinia pe |
| 24         | 189.5 | 13.7          | 605    | 2 Q6K9D3     | Q6k9d3 oryza sativ |
| 25         | 188.5 | 13.6          | 922    | 2 Q9L964     | Q9l964 moraxella c |
| 26         | 188.5 | 13.6          | 1011   | 2 Q6Z0A9     | Q6z0a9 oryza sativ |
| 27         | 187.5 | 13.5          | 784    | 2 Q9LS79     | Q9ls79 arabidopsis |
| 28         | 186.5 | 13.5          | 465    | 2 Q8I4Z3     | Q8i4z3 plasmodium  |
| 29         | 186   | 13.4          | 641    | 2 Q9LDG0     | Q9ldg0 oryza sativ |
| 30         | 186   | 13.4          | 645    | 2 Q9SVN1     | Q9svn1 arabidopsis |
| 31         | 185   | 13.4          | 1012   | 2 Q7XP13     | Q7xp13 oryza sativ |

32 184.5 13.3 836 2 Q9SCT4 Q9sct4 arabidopsis  
33 183.5 13.2 380 2 Q6Z3T9 Q6z3t9 oryza sativ  
34 183.5 13.2 1051 2 Q67IT2 Q67it2 oryza sativ  
35 182 13.1 941 2 Q9XD54 Q9xd54 moraxella c  
36 181.5 13.1 718 2 Q6ET59 Q6et59 oryza sativ  
37 181 13.1 653 2 Q8LKV9 Q8lkv9 aegilops ta  
38 181 13.1 883 2 Q9M9X1 Q9m9x1 arabidopsis  
39 181 13.1 1065 2 Q9LGI5 Q9lgi5 oryza sativ  
40 180.5 13.0 703 2 Q6R2J8 Q6r2j8 arabidopsis  
41 180 13.0 998 2 Q8LJI9 Q8lji9 oryza sativ  
42 180 13.0 1047 2 Q6K2I3 Q6k2i3 oryza sativ  
43 180 13.0 1080 2 Q69P46 Q69p46 oryza sativ  
44 179.5 13.0 983 2 Q8L7L6 Q8l7l6 arabidopsis  
45 179.5 13.0 983 2 Q9ZUK3 Q9zuk3 arabidopsis

#### ALIGNMENTS

RESULT 1  
Q9M3W4 PRELIMINARY; PRT; 118 AA.  
ID Q9M3W4  
AC Q9M3W4;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Ice recrystallisation inhibition protein (Fragment).  
OS Lolium perenne (Perennial ryegrass).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Poae; Lolium.  
OX NCBI\_TaxID=4522;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sidebottom C.M.;  
RL Thesis (1999), University of York.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Sidebottom C.M.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ277399; CAB87814.1; -  
DR PDB; 1I3B; Model; A=1-118.  
FT CHAIN <1 118  
FT NON\_TER 1  
SQ SEQUENCE 118 AA; 11766 MW; 0362C1B4F98AE9F8 CRC64;  
Query Match 38.3%; Score 530; DB 2; Length 118;  
Best Local Similarity 85.5%; Pred. No. 3.5e-32;  
Matches 100; Conservative 10; Mismatches 7; Indels 0; Gaps 0;  
QY 150 DEEPNTISGNNVSGSGNNVSGNDNTVTVSGNNHVSNNNTVTVSGNDNTVTVSGNNHVS 209  
Db 1 DQQNTISGNNVTVSGSKVNLGNDNTVSGDNNVSGNNNTVTVSGNDNTVTVSGNNHVS 60  
QY 210 GTKGHVTDNNNVSGNDNNVSGSFHTVSGEHTVSGNNNTVSGNNHVSNNKVVTD 266  
Db 61 GTNHVTDNNNVSGNDNNVSGSFHTVSGEHTVSGNNNTVSGNNHVSNNKVVTD 117  
RESULT 2  
Q8H3W8 PRELIMINARY; PRT; 1010 AA.  
ID Q8H3W8  
AC Q8H3W8;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DE Putative phytoosulfokine receptor.  
GN Name=P0585H1.109;  
OS Eryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;







```
Query Match 14.7%; Score 204; DB 1; Length 1021;
Best Local Similarity 24.6%; Pred. No. 1.2e-06;
Matches 81; Conservative 56; Mismatches 94; Indels 98; Gaps 14;

QY 10 MLVFLGPIQ--VAGATWSCHDDHAL-----RGLAENLSGKGAVALRAANS-----57
DB 8 VILLVGFQVIVVNSQNTNSNDKALEGFMRLGLESSIDG-----WKWNESS 58
QY 58 -GASCCSWEVGCEA-----SGRVVALRLPKRGIGGIIPSSIGELDLRLVLDLS 106
DB 59 FSNCCDWGISCKSSVSLGDDWESGRVVELEGRLKLSGKLSSEVAKLDQLKVLNT 118
QY 107 GNSLVGEVPSKL-----QIRLKSITTDQSGLMGSIINMLHVSSRRRLTDEEPNTISG----158
DB 119 HNSLSGSIASLNLNLEVLDSNDFS---GLTFPSLINLPLSLRVLVVYENSFHGLIPA 175
QY 159 --TNN-----SVGSGSNVNVSGNDNTV--VSGNNHVS-----SNNTVV 194
DB 176 SLNCNLPRIRIEDIAMNYPDGSIPVIGNCSSVEYLGASNNLSGSIPOELFQLSNLVL 235
QY 195 TGSNDTVVSGNHVSGTKHIVTNNNVVSGNDNNVSGSFHV-----SGEHNTVS 244
DB 236 ALQNNRLSGA-----LSKLGKLNGLDLISSNKFSGKIPDVPLELNKLVYFSAQNLFN 291
QY 245 GS-----NNTVSGSNHI 256
DB 292 GEMPSLSNSRSISLLSRNNTLSGGIYL 320

RESULT 8
Q66QA4 PRELIMINARY; PRT; 1065 AA.
ID Q66QA4
AC Q66QA4;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DE Putative leucine-rich repeat receptor-like kinase.
OS Oryza sativa (indica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39946;
RN [1]
RN SEQUENCE FROM N.A.
RP He G.M., Li K.G., Yang J.S.;
RT "Oryza sativa (indica cultivar-group) chromosome 2 genomic sequence.";
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP He G.M., Li K.G., Yang J.S.;
RT TISSUE=Leaf;
RT "Isolation and expression pattern of putative LRR (leucine-rich repeat) receptor-like kinase in rice.";
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY730046; AAU12611.1;
DR EMBL; AY714494; AAU12603.1;
DR GO; GO:0016301; F-kinase activity; IEA.
DR GO; GO:0004872; F-receptor activity; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003885; LRR_cyst.
DR InterPro; IPR007090; LRR_plant.
DR InterPro; IPR003591; LRR_typ.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00560; LRR_1; 20.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00365; LRR_SD22; 6.

DR SMART; SM00369; LRR_TYP; 8.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYKIC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase; Receptor.
SQ SEQUENCE 1065 AA; 116827 MW; CFE570663C05BFOA CRC64;

Query Match 14.5%; Score 201; DB 2; Length 1065;
Best Local Similarity 24.6%; Pred. No. 2.1e-06;
Matches 87; Conservative 54; Mismatches 98; Indels 114; Gaps 16;

QY 7 KCMLAVFLGF-----ILOVAGATWSCHDDHALRGLAENLSGKGAVALRAA 55
DB 12 KCSNRCFTDFLRFNLAFAALLLLSLASPTS-SCTEQERSLLQLFLSSLSDDGLAV-SW 69
QY 56 WSGASCCSWEVGCEAASGRVVALRLPKRGIGGIIPSSIGELDLRLVLDLSGNSLVCEVP 115
DB 70 WNAADCCKWEGVTC-SADGTVTVDVSLASKGLEGRISPSLGNLTGLRLRLNLSHNSLSGGLP 128
QY 116 KSLQIRLKSITTDQSGLMGSIINML-----LHVSSRRRLTDEEPNT-- 155
DB 129 -----LELMASSITVLDISFNLLKEETHELPSSTPARPLOVINISSNLTQGFPSATW 182
QY 156 -----ISGTNNS-VGSGSNVNVSGNDN-TVSGNNHVS-----NNTVVTG 196
DB 183 EMMKNLVMLNASNNSPTGOIPSNFCSRSPSLTVLALCYNHLNGSIPPFGNCLKLRVLKA 242
QY 197 SDNTVVS-----NHVVSQT-----KHIVT-----DNNV----- 221
DB 243 GHNNLSGNLPGDLFNATSLLEYLFPNNELNGVINGTLIVNRLNLTDLDEGNNINGRPD 302
QY 222 -----VSGNDNNVSGSF-----H--TVSGEHTVSGSNNTVSGSN 254
DB 303 SIGQLKRLQDLHLGDNNISGELPSALSNCNTHLITINLKNNFNSGNLSNVNFSN 355

RESULT 9
Q9C9H6 PRELIMINARY; PRT; 784 AA.
ID Q9C9H6
AC Q9C9H6;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Putative disease resistance protein; 69620-67266.
GN Name-P26A9.23;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RP Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Renning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC016163; AAG51813.1;
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007090; LRR_plant.
DR Pfam; PF00560; LRR_1; 20.
DR PRINTS; PR00019; LEURICHRPT.
DR SQ SEQUENCE 784 AA; 88730 MW; 8770246421192623 CRC64;

Query Match 14.4%; Score 199.5; DB 2; Length 784;
Best Local Similarity 24.2%; Pred. No. 1.9e-06;
Matches 80; Conservative 51; Mismatches 80; Indels 119; Gaps 15;
```

[illegible]

```

RESULT 10
Q940M7 PRELIMINARY; PRT; 751 AA.
ID AC Q940M7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AT5967280/K3G17.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowers L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.
RL Submitted (AUG-2001) to the ENMBL/GenBank/DBJ databases.
RL ENMBL; AY054256; AAL06915.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; P:protein kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007090; LRR_plant.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00560; LRR_1; 5.
DR PRINTS; PR00019; LEURICHRPT.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
SQ SEQUENCE 751 AA; 82086 MW; E6ADAB4370F49689 CRC64;

```

|                       |        |  |                           |                    |
|-----------------------|--------|--|---------------------------|--------------------|
| Query Match           | 14.3%  | Score 197.5;   | DB 2;                     | Length 751;        |
| Best Local Similarity | 28.6%; | pred. No. 2.5e-06;                                       |                           |                    |
| Matches               | 67;    | Conservative 34;   | Mismatches 78;            | Indels 55; Gaps 9; |
| QY                    | 62     | CSWEGVGTCTAGRWALRKRGLGII                                 | PSSIGELDLRLYLDLSNVLGVPKSL | --- 118            |
| Db                    | 64     | CSWRGVTCDASRHVTVLSLPSSNLTGLPSNLGSLNSQLRDLNNSNGSFPVSLNA   | 123                       |                    |
| QY                    | 119    | -QIRLKSITTT-----SQSLGMSINMLAHVSRRLTDEEPTNTISGTNNVSGSGNNW | 171                       |                    |

|    |     |   |     |
|----|-----|---|-----|
| Db | 124 | TELRLDLSNDHISGALPASFCALSNLQVLNLSDNSFVGEIPLNTL-GWNRNL----- | 174 |
| Qy | 172 | SGNDNTVTVSGNNHVS-----SNNVTGSDNTTVGS--NHVYSGTKHIVTDNNNVVS  | 223 |
| Db | 175 | -----TELQKNVLSGGIPGGFKSTVTELDLSSNLIKSLPSHFRGNRLRFPNYSNRIS | 229 |
| Qy | 224 | G-----NDNNVSGSFHTVSG-----EHTVTVSGSNNTVSGSNH               | 255 |
| Db | 230 | GEIPSGFADEIPEDATVDLSFNQLTGQIPGPRVLDNQBSNFSFG-NPGLCGSDH    | 282 |

  

|           |  |   |              |
|-----------|--|---|--------------|
| RESULT 11 |  |   |              |
| Q9FGQ5    |  |   |              |
| ID        | Q9FGQ5   | PRELIMINARY;                            | PRT; 751 AA. |
| AC        | Q9FGQ5;  |   |              |
| DT        | 01-MAR-2001  | (TrEMBLrel. 16, Created)                |              |
| DT        | 01-MAR-2001  | (TrEMBLrel. 16, Last sequence update)   |              |
| DT        | 05-JUL-2004  | (TrEMBLrel. 27, Last annotation update) |              |
| DE        | Receptor-like protein kinase (At5g67280/K3G17.4).                    |   |              |
| OS        | Arabidopsis thaliana (Mouse-ear cress).                              |   |              |
| OC        | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;   |   |              |
| OC        | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; |   |              |
| OC        | eurosid II; Brassicales; Brassicaceae; Arabidopsids.                 |   |              |
| OX        | NCBI_taxID=3702;   |   |              |
| RN        | [1]  |   |              |
| RP        | SEQUENCE FROM N.A.   |   |              |
| RA        | Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,    |   |              |
| RA        | Tabata S.;   |   |              |
| RL        | Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.              |   |              |
| RP        | [2]  |   |              |
| RP        | SEQUENCE FROM N.A.   |   |              |
| RA        | Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,       |   |              |
| RA        | Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,        |   |              |
| RA        | Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., |   |              |
| RA        | Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Palm C.J.,     |   |              |
| RA        | Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,  |   |              |
| RA        | Toriiumi M., Wallender E.K., Wong C., Wu H.C., Yamada K., Yu G.,     |   |              |
| RA        | Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;          |   |              |
| RL        | Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.              |   |              |
| DR        | EMBL; AB025614; BAB09647.1; -.                                       |   |              |
| DR        | EMBL; BT001094; AAN46893.1; -.                                       |   |              |
| DR        | GO; GO:0005524; F:ATP binding; IEA.                                  |   |              |
| DR        | GO; GO:0004672; F:protein kinase activity; IEA.                      |   |              |
| DR        | GO; GO:0004872; F:receptor activity; IEA                             |   |              |
| DR        | GO; GO:0006468; P:protein amino acid phosphorylation; IEA.           |   |              |
| DR        | InterPro; IPR011009; Kinase_like.                                    |   |              |
| DR        | InterPro; IPR001611; LRR.  |   |              |
| DR        | InterPro; IPR007090; LRR_plant.                                      |   |              |
| DR        | InterPro; IPR000719; Prot_kinase.                                    |   |              |
| DR        | Pfam; PF00560; LRR_1; 5.   |   |              |
| DR        | PRINTS; PR00019; LEURICHRPT.   |   |              |
| DR        | PRODOM; PD000001; Prot_kinase; 1.                                    |   |              |
| DR        | PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.                             |   |              |
| DR        | Kinase Receptor.   |   |              |
| Qy        | SEQUENCE 751 AA; 82062 MW; 180302F2B7E9FF966 CRC64;                  |   |              |

|                       |        |   |       |                                   |
|-----------------------|--------|---|-------|-----------------------------------|
| Query Match           | 14.3%; | Score 197.5;  | DB 2; | Length 751;                       |
| Best Local Similarity | 28.6%; | Pred. No. 2.5e-06;  |       |                                   |
| Matches               | 67;    | Conservative  | 34;   | Mismatches 78; Indels 55; Gaps 9; |
| Qy                    | 62     | CSWEGVCETASGRVVALRLPKRLGIIIPSSIGELDHLRYLDLSGNLSIGVEVPKSL---   | 116   |                                   |
| Db                    | 64     | CSWRGVTCDASSRHVTVLSPSSNLTLPSNLGSLNSLQRLDLSNNSINGSFVPVSLNA     | 123   |                                   |
| Qy                    | 119    | -QIRLKSLLTVD-----SQSLGMSGIMMLHIVSSRRRLDDEPNTISCTGNNSVSGSGSNVV | 171   |                                   |
| Db                    | 124    | TELRFDLSDNHISGALPASFGALSNLQVLNLSNDSFVGELPNTL-GWNRNL-----      | 174   |                                   |
| Qy                    | 172    | SGNDNTVVGSGNNHVG-----SNNTVTGSNDTVVGS---NHVVGSKHIVTDNNNVVS     | 223   |                                   |
| Db                    | 175    | -----TEISLQRYLGGIPGGKSTEYLDLSNLIGKSPSHFRGNRLRYFNAYSYNRIS      | 229   |                                   |
| Qy                    | 224    | G-----NDNNVVGSGFHTVSG-----EHTVSGSNNTVSGSNH                    | 255   |                                   |

```
Db 230 GEIPSGFADEIPDATVDLSFNLTQIGQIFRVLNDQESNFSFG-NPGLGSDH 282
RESULT 12
Q69KC3 PRELIMINARY; PRT; 1102 AA.
AC Q69KC3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B1047H05.16.
GN Name=B1047H05.16;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
clone:B1047H05.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AF005966; BAD34184.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007090; LRR plant.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00560; LRR_1; 24.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PD00019; LEURICHRPT.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00369; LRR_TYP; 7.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1102 AA; 119984 MW; B960CDE45175C68E CRC64;

Query Match 14.2%; Score 196; DB 2; Length 1102;
Best Local Similarity 24.3%; Pred. No. 5.2e-06;
Matches 84; Conservative 47; Mismatches 109; Indels 106; Gaps 14;

QY 13 VFLGFIQVAGATSWCHHDDLHALRGLAENLSGKGAVALRAAWGASC--CSWEGVGCE 70
DB 19 IFCSIATINCTN-----DDROALLCFKSQLSGPS--RVLSWSNTSLNFCNWDGVTCS 70
QY 71 TAS-GRVALRLPKRG-----LGGIIPSSIGELDLRLYDL 105
DB 71 SRSPFVIAIDLSSEGITGTISPCIANLTSMTLQIENSLHGSIPPKGLLRKLRNL 130
QY 106 SGNSLGVCEPKSL-----QIRKSLTTDS-----QSLG----- 133
DB 131 SMNSLEGNIPQLSSYSQIBILDSSNFSQGAIPASLGKCIHLQDINLSRNQLQRISSA 190
QY 134 MGSINML--LHVSRRRLDEPNITIS-----GTNNSVSGSNVSGNDNTVSG 181
DB 191 FGNLSKLOALVLTNSRLTDEIPPLSGSSFSFLRYVDLGNNDITGSIPESLANSSSLQVRL 250
QY 182 NNNHVS-----SNNTVVTGSDNTVVGSHVSGTKHIVTDNNNVSGNDNVSGS 232
```

```
Db 251 MNNLSGVEPKSLFNTSSITAIQLONSFVSGIPAAMSSPI-----KYISLRDNCISGT 306
QY 233 FHTVSGEHTVSG-----GNN-----TVSGSNHIVSGSNKV 264
DB 307 IPESLGHTRLTILMTSVNNLSGLVPPSLFNISLTFILAGNNSLV 352

RESULT 13
Q84NG8 PRELIMINARY; PRT; 1023 AA.
AC Q84NG8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative receptor kinase.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
SEQUENCE FROM N.A.
RA Gu Y.Q., Anderson O.D., Londeore C.F., Kong X., Chibbar R.N.,
RA Lazo G.R.;
RT "Structural organization of the barley D-hordein locus in comparison
with its orthologous regions of wheat genomes.";
RL Genome 46:1084-1097(2003).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AY268139; AAP31049.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007090; LRR plant.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00560; LRR_1; 17.
DR PRINTS; PD00019; LEURICHRPT.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW Transferase.
SQ SEQUENCE 1023 AA; 110187 MW; 09EB62D3C41141C6 CRC64;

Query Match 14.1%; Score 195.5; DB 2; Length 1023;
Best Local Similarity 30.3%; Pred. No. 5.2e-06;
Matches 84; Conservative 37; Mismatches 113; Indels 43; Gaps 13;

QY 2 PEYAKCCMLLVFLGFIQVAGATSWCHHDDLHALRGLAENLSGKGAVALRAAW----S 57
DB 5 PWLLRLIALITTTALLLPSTSSSVSTAH-DLPALLSP-KSLITKDPGLASSWTNGS 62
QY 58 GASCCSWEGVGCEA-SGRVALRLPKRGIGIIPSSIGELDLRLYDLISGNSLVEGPK 116
DB 63 THGFCSWTGVCESSAHPGHVKALRQLGLSGTISPFGLNLSRLRALDLSGNKLOQIPS 122
QY 117 SL-----QIRKSLTTDS-----LNGSINMLLHVSSRRRLDEPNITISCTNNSVSGSN 168
DB 123 SIGNCFALRTLNLSVNSLGAIPPMGNLSKLL-----VLSVKNDISGTITPSFAGLA 176
QY 169 NVVSGNDNTVSGNNHVS-----SNNTVVTG-----SDNTVVGSHVSGTKHIVTDNN 219
DB 177 TV-----AVFSVARNHVGQVPPWLCNLTALDNLNADN--IMSGHVPALSKLINRS 228
QY 220 NVVSGNDNVSGSFHTVSGEHTVSGSN---NTVSGS 253
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2005, 20:18:45 ; Search time 16 Seconds  
(without alignments)  
1605.617 Million cell updates/sec

Title: US-10-657-852A-15  
Perfect score: 1385  
Sequence: 1 MPEYNAKCCMLLVFLGLQ.....NTVSGSNHIVSGSNKVVTG 267.  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues 283416  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 190.5 | 13.8        | 1008   | 2 D84434 | probable receptor- |
| 2          | 190   | 13.7        | 622    | 2 AF0169 | probable exported  |
| 3          | 186   | 13.4        | 645    | 2 T05251 | probable disease r |
| 4          | 184.5 | 13.3        | 836    | 2 T46070 | hypothetical prote |
| 5          | 179.5 | 13.0        | 983    | 2 G84524 | probable disease r |
| 6          | 176.5 | 12.7        | 743    | 2 C84633 | probable protein k |
| 7          | 173.5 | 12.5        | 590    | 2 B86440 | hypothetical prote |
| 8          | 169   | 12.2        | 420    | 2 AE1857 | hypothetical prote |
| 9          | 168.5 | 12.2        | 1095   | 2 G96746 | probable disease r |
| 10         | 168   | 12.1        | 1019   | 2 C96519 | probable disease r |
| 11         | 164.5 | 11.9        | 716    | 2 B84421 | probable receptor- |
| 12         | 164   | 11.8        | 766    | 2 B85440 | receptor kinase-li |
| 13         | 162.5 | 11.7        | 1232   | 2 T05322 | hypothetical prote |
| 14         | 162   | 11.7        | 773    | 2 T00502 | probable receptor- |
| 15         | 161.5 | 11.7        | 905    | 2 T00475 | probable disease r |
| 16         | 161   | 11.6        | 719    | 2 T47727 | hypothetical prote |
| 17         | 161   | 11.6        | 907    | 2 A86460 | 99.9K hypothetical |
| 18         | 159.5 | 11.5        | 768    | 2 T17462 | disease resistance |
| 19         | 158.5 | 11.4        | 658    | 2 B84664 | probable receptor- |
| 20         | 158   | 11.4        | 478    | 2 H86459 | hypothetical prote |
| 21         | 157.5 | 11.4        | 996    | 2 T10725 | protein kinase Xa2 |
| 22         | 157   | 11.3        | 241    | 2 S61925 | cyst wall protein  |
| 23         | 156.5 | 11.3        | 702    | 2 A86383 | 76.4K protein kin  |
| 24         | 156.5 | 11.3        | 876    | 2 D96558 | probable protein k |
| 25         | 156   | 11.3        | 371    | 2 T49908 | hypothetical prote |
| 26         | 156   | 11.3        | 638    | 2 T05606 | protein kinase hom |
| 27         | 155.5 | 11.2        | 942    | 1 JQ1674 | protein kinase TMK |
| 28         | 155   | 11.2        | 729    | 2 F86308 | Similar to disease |
| 29         | 154.5 | 11.2        | 1143   | 2 T10636 | hypothetical prote |

|    |       |      |     |          |                    |
|----|-------|------|-----|----------|--------------------|
| 30 | 154   | 11.1 | 720 | 2 T02361 | hypothetical prote |
| 31 | 153   | 11.0 | 668 | 2 T05257 | probable disease r |
| 32 | 152.5 | 11.0 | 646 | 2 B84852 | probable receptor- |
| 33 | 152.5 | 11.0 | 655 | 2 G96524 | protein T1N15.9 [i |
| 34 | 152.5 | 11.0 | 838 | 2 A96557 | probable receptor  |
| 35 | 152.5 | 11.0 | 980 | 2 H84632 | probable receptor- |
| 36 | 151.5 | 10.9 | 480 | 2 T00971 | probable disease r |
| 37 | 151   | 10.9 | 693 | 2 F96557 | probable protein k |
| 38 | 151   | 10.9 | 863 | 2 F85343 | hypothetical prote |
| 39 | 151   | 10.9 | 981 | 2 T50851 | receptor protein k |
| 40 | 150.5 | 10.9 | 221 | 2 T07079 | leucine-rich repea |
| 41 | 150.5 | 10.9 | 672 | 2 B84782 | probable receptor- |
| 42 | 150.5 | 10.9 | 813 | 2 T04313 | protein kinase Xa2 |
| 43 | 149.5 | 10.8 | 286 | 2 C61615 | sericin MG-2 - gre |
| 44 | 149.5 | 10.8 | 645 | 2 E96631 | probable receptor  |
| 45 | 149   | 10.8 | 976 | 2 B84659 | probable receptor- |

ALIGNMENTS

RESULT 1

D84434  
probable receptor-like protein kinase [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C;Accession: D84434  
M.; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: D84434  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-1008 <STO>  
A;Cross-references: GB:AE002093; NID:g6598459; PIDN:AAC78507.2; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2g02220  
A;Map position: 2

|                       |        |  |   |            |     |        |       |
|-----------------------|--------|--|---|------------|-----|--------|-------|
| Query Match           | 13.8%; | Score  | 190.5;  | DB         | 2;  | Length | 1008; |
| Best Local Similarity | 27.5%; | Pred. No.  | 3e-06;  |            |     |        |       |
| Matches               | 68;    | Conservative   | 45;   | Mismatches | 93; | Indels | 41;   |
| Gaps                  | 12;    |  |   |            |     |        |       |
| QY                    | 9      | CMLLVFLG----   | FILQVAGATSWSCHDDHLALGLAENLSKGAVRLRAAWSGASCCSW | 64         |     |        |       |
| Db                    | 7      | CVIVFLTELLCFFYSSSQTTSRCHPHDLEALRDPFAHLEPKPDGWINSS-SSTDCCNW   | 65  |            |     |        |       |
| QY                    | 65     | EGVGCT-ASGRVVALRLPKRGLGGIIPSSIGELDHLRYLDLSGSLVGEVPKSLQIRLK   | 123   |            |     |        |       |
| Db                    | 66     | TCITCNSNTGVRVIRLEGLNGKLSGKLSKLSKLDLRLVNLNLSRNFIKDSIPLSI-FNLK | 124   |            |     |        |       |
| QY                    | 124    | SITT---DSQSLGMSINMLLVSSRRITDEPNTISGT-----NNS-----VGSGSNV     | 170   |            |     |        |       |
| Db                    | 125    | NLQTLDSLNDLS-GGIPTSINLPALQSPDLSSNKFNGSLPSHICHNSQTQIRVVLAVNY  | 183   |            |     |        |       |
| QY                    | 171    | VSGNDTVVSGNNHVGSGNNTVVTGSDNTVVGSHVSGTKHIVTD-----NNNVVSGN     | 225   |            |     |        |       |
| Db                    | 184    | FAG-----NFTSGFKCVLL--EHLCLGMNDL---TGNIPEDLFLKRLNLLGIQ        | 228   |            |     |        |       |
| QY                    | 226    | DNNVSGS  | 232   |            |     |        |       |
| Db                    | 229    | ENRLSGS  | 235   |            |     |        |       |

RESULT 2

AF0169  
probable exported protein YP01388 [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AF0169

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell, Nature 413, 523-527, 2001  
 A>Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
 A;Reference number: AB0001; MUID:21470413; PMID:11586360  
 A;Accession: AF0169  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-622 <KUR>  
 A;Cross-references: UNIPROT:Q8ZGB5; GB:AL590842; PIDN:CAC90217.1; PID:gl5979437; GSPDB:Q  
 C;Genetics:  
 A;Gene: YP01388

Query Match 13.7%; Score 190; DB 2; Length 622;  
 Best Local Similarity 26.3%; Pred. No. 1.8e-06;  
 Matches 68; Conservative 49; Mismatches 72; Indels 70; Gaps 13;  
 QY 43 NLSGKGVRLRAAMSGASCCEWGVGCTASGRVVALRPLKRGGLGPIPSIGELDLHLYR 102  
 DB 41 NNNGNGTINIPDASNNDIHILTLGNEQ-----LGGF-----SNHL-- 77  
 QY 103 LDLSGNSLVGEVPSKQLRLKSLTDSQ-----SLGMG-----SINMLLHVSS 145  
 DB 78 IDSHNTTIDGQSNL-----VSSDGNMISALSLDGLFYCAQNTNLINSNNLLIVTQ 131  
 QY 146 RRT-LDEPNTISGNTNNSVSGSNVNSGNDNTVVS-----GNNHVSNNNTV 193  
 DB 132 GSTIDSNTVSGISNNLIESNNII-GNENSCYSDPASPAGMVCVDNQTLLGSDNNT 190  
 QY 194 VTGSDNTVVGNNH--VVSQTKHIVTDNNNVVSGNDNVSGSFHTVS-GEHNTVSGNNNT 249  
 DB 191 ITGALGNLHSHNDIIASSVNNLMDTHNIAG-----GHYTIISGCGNDIFGSENN 244  
 QY 250 VSGSNH1-VSGSNKVVTDG 267  
 DB 245 VTDSDANINGSNNVVDG 263

RESULT 3  
 T05251  
 Probable disease resistance protein F18A5.210 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C;Accession: T05251  
 R;Bevan, M.; Weber, N.; Gruening, D.; Schmidheini, T.; Bancroft, I.; Mewes, H.W.; Maye  
 submitted to the Protein Sequence Database, February 1999  
 A;Reference number: Z15405  
 A;Accession: T05251  
 A;Molecule type: DNA  
 A;Residues: 1-645 <BEV>  
 A;Cross-references: UNIPROT:Q9SVN1; EMBL:AL035528  
 A;Experimental source: cultivar Columbia; BAC clone F18A5  
 C;Genetics:  
 A;Map position: 4  
 A;Introns: 216/1; 541/1  
 A;Note: F18A5.210

Query Match 13.4%; Score 186; DB 2; Length 645;  
 Best Local Similarity 22.1%; Pred. No. 3.8e-06;  
 Matches 82; Conservative 46; Mismatches 115; Indels 128; Gaps 14;  
 QY 9 CMLLVFLGFILOVAGATSWSCCHDDHLALRGAE-----NLSGKGVRLRAAW-SGAS 60  
 DB 7 CLILSLNSKLVLAHVHKLQRODKNALBEFKNEFYVHEFNNGINGVGVKTEKWRNTD 66  
 QY 61 CCSWEGVGCETASGRVVALRL-----PKR-----GLGGLIIPSSI 94  
 DB 67 CCSWDGISCDPKTKGVVELDLNNSFLNGLFLYDSSLFLQLHLNLDLSNNFSGILPDSI 126  
 QY 95 GELDLRLYLDLSGNSLVGEVPSKQLRLKSLTDSQSL-----GMGSINML--LHV 143  
 DB 127 GSKLYRLVSLGDCNLFQKIPSSLG-NLTYLTNLDLSVNDPTGELPDSMGHLNKLTELHL 185

QY 144 SSRTLDPEPNTISGTH--NSVSGSN-----NVVSGN-----DN 176  
 DB 186 GSAKSLGNFFSMLNLSELTLIDLSNQFGNQTLYLDISANKIGQVFPQWLWSP 245  
 QY 177 TVSGNNHVSQ-----SNNVTVTGSDNTVVS 204  
 DB 246 QYVNISQNSFGPGPADVIQRCGELMLDISNTFQDPFPLPNSTIFLGSNRRSGE 305  
 QY 205 -----NHVVGSGKHIVTDNN-----NVVSGNDNVVSGSF--HTVSGEHN 241  
 DB 306 IPKTKICKLVSLDTLVNNSNNFNGSIPRCFEKFNFTLSVLRNNSLSEPPESISDHLR 365  
 QY 242 TVSGSNNTVSG 252  
 DB 366 SLDVGRNRLSG 376

RESULT 4  
 T46070  
 Hypothetical protein T18N14.120 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 16-Aug-2004  
 C;Accession: T46070  
 R;Delseny, M.; Berger, C.; Cooke, R.; Grellier, P.; Laudie, M.; Mewes, H.W.; Lemcke, K.;  
 submitted to the Protein Sequence Database, December 1999  
 A;Reference number: Z23013  
 A;Accession: T46070  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-836 <DEL>  
 A;Cross-references: UNIPROT:Q9SCT4; EMBL:AJ132968  
 A;Experimental source: cultivar Columbia; BAC clone T18N14  
 C;Genetics:  
 A;Map position: 3  
 A;Introns: 625/1  
 A;Note: T18N14.120  
 C;Superfamily: protein kinase homology

Query Match 13.3%; Score 184.5; DB 2; Length 836;  
 Best Local Similarity 25.7%; Pred. No. 6.7e-06;  
 Matches 82; Conservative 48; Mismatches 100; Indels 89; Gaps 15;  
 QY 10 MLLVFLGFILOV-----AGATSW-----SCHDDHLALRGAE--NLSGKGVRLRAAWSG 58  
 DB 22 LLILCFLLFSAQAVAGGGGSHSDGIVVTOANYOALQAIKHELIDFTG-----VLKSWNN 76  
 QY 59 AS-----CCSWEGVGCETASGRVVALRPLKRGGLGPIPSIGELDLHLYLDLSGNSLVGEV 114  
 DB 77 SASQVCSGWAGIKC--LRQGVVAIQLPWKGLGTISEKIGQLSLRKLSHNNVIAVS 134  
 QY 115 PKSLQIRLKS-----TTDSQSLGMSINMLHVSRRRTDEPNTISGT----- 159  
 DB 135 PRSLGY-LKSLRGVYLFNNRLSGSIPVSLGNCPL-----QNLDLSNQLTGAIPPSLT 187  
 QY 160 -----NNSVSGSNVNSGNDNVVSGSNVNSVSGS--NNTVVTGS-----D 198  
 DB 188 ESTRLYRLNLSFNSLSGFLPVARSYTLFLDLQHNLSGSDPFFVNGSHPLKTLNLD 247  
 QY 199 NTVVSGNHVVSQTKHIVTDNNNVVSGNDNVVSGSFHTVSG-----EHNVTSGS-- 246  
 DB 248 HNRFGAVPVSLCKHSLLEB---VSISHNQLSGSIPRECGLPHLQSLDLSYNSINGTIP 304  
 QY 247 -----NNTVS-----GSNHI 256  
 DB 305 DSFNSLSLVSLENNHL 323

RESULT 5  
 G84524  
 Probable disease resistance protein [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

| Query Match           | 12.7%           | Score 176.5       | DB 2      | Length 743 |
|-----------------------|-----------------|-------------------|-----------|------------|
| Best Local Similarity | 26.7%           | Pred. NO. 2.3e-05 |           |            |
| Matches 63            | Conservative 35 | Mismatches 81     | Indels 57 | Gaps 10    |

|                       |       |   |                |                    |
|-----------------------|-------|---|----------------|--------------------|
| Query Match           | 12.5% | Score 173.5;  | DB 2;          | Length 590;        |
| Best Local Similarity | 24.7% | Pred. No. 2.8e-05;  |                |                    |
| Matches               | 70;   | Conservative 44;  | Mismatches 90; | Indels 79; Gaps 13 |
| Qy                    | 5     | MAKCCMLLVFLGFLQVAGATSWCHDDDLHALRGLAENLRSKGVRLRAA-----         | 55             |                    |
| Db                    | 6     | MKSCCSWLLLTSLCSLSNESQ-----AISPDEALLSPFNATVTRSDSFTH            | 51             |                    |
| Qy                    | 56    | W--SGASCSCWEGVGCETASGRVVALRLPKRGLGGIIPSSIGELDLHLYRLDLSGNSLWG  | 112            |                    |
| Db                    | 52    | QWRPEDPDCNWNWGVTCDAKTRVITLNLTYHKIMGPLPPDGKUDHLRLMLHNNALYG     | 111            |                    |
| Qy                    | 113   | EVPKSL-----QIRLKS-LTTDSQSLGMSGINMLLHVSRRTLDEBPNTISG-----      | 158            |                    |
| Db                    | 112   | AIPTALGNCTALBEIHLQSNYFTGPIPAEMGDLPCL-----QKLDMSNTLSGPIASL     | 165            |                    |
| Qy                    | 159   | -----TNNSYSGSNVNVVSGN--DNTVTVSGNNNHVSGSNNTVTVTGSQNTVTVSGNHVSG | 210            |                    |
| Db                    | 166   | GQLKLSNFNV--SNNFLVQGFPSDGLVSG-----FSKNSPFIGNLNLG- 207         |                |                    |
| Qy                    | 211   | TKHIVTNNNVVSGDNNVSGSFHTVSGEHTVSGSNNTVSGS                      | 253            |                    |
| Db                    | 208   | -KHV-----DVVCCDDSG-NPSSHQSQGNQKQSGKLLISAS                     | 243            |                    |

AE1857

C;Species: Nostoc sp. PCC 7120

C/Species: Nostoc sp. PCC 7120  
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120



M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: H84421  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-716 <STO>  
A;Cross-references: UNIPROT:Q9ZU46; GB:AB002093; NID:g4262228; PIDN:AAD14521.1; GSPDB:GN  
C;Genetics:  
A;Gene: At2g01210  
A;Map position: 2

Query Match 11.9%; Score 164.5; DB 2; Length 716;  
Best Local Similarity 27.4%; Pred. No. 0.00017;  
Matches 63; Conservative 34; Mismatches 78; Indels 55; Gaps 9;

QY 10 MLIVFLGFILOVAGATSWCHDDHALRGLAENLSKGAVRLRAAW--SGASCCSWGV 67  
Db 1 MLASLIIFVALLCNVTIVISGLNDEGFALLTFKQSVHDDPTGSLN--WNNSDENACSWNGV 59

QY 68 GCBTASGRVVALRPRKGLGGIIPSSIGELDLRLYLDLSCNSLVGEVP-----KSL 118  
Db 60 TCKEL--RVVSLIPRNLVYSLPSSIGFLSSRLHLNLRNRYGSLPTQLFLHQLQLSL 117

QY 119 QI-----RLKSLTT--DSQSLGMSINN--LHVSSRRRLDEEPTISGT-- 159  
Db 118 VLYGNSFDGSLSEIGKLLQLDLSONLFNSLPLSLIQCNRKLTLDVSRNLSGLPL 177

QY 160 -----NNSVSGSNVNVSGNDNTVSGNHNHVS 189  
Db 178 DGFSAFVSLKLDLAPNQFNQSPIDIGNL--SNLQGTADFSHNHTGS 225

RESULT 12  
B85440

receptor kinase-like protein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Aug-2004  
C;Accession: B85440

R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999

A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: B85440  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-766 <STO>  
A;Cross-references: UNIPROT:O23161; GB:NC\_001268; NID:g7270708; PIDN:CAB80391.1; GSPDB:G  
C;Genetics:  
A;Gene: At4g37250  
A;Map position: 4  
C;Superfamily: protein kinase homology

Query Match 11.8%; Score 164; DB 2; Length 766;  
Best Local Similarity 28.5%; Pred. No. 0.0002;  
Matches 63; Conservative 35; Mismatches 67; Indels 56; Gaps 9;

QY 60 SCSSWGVGCEATAGRVALLRPRKGLGGIIPSSIGELDLRLYLDLSCNSLVGEVPKSL- 118  
Db 51 SPCSWRGISCNNDG--KVLTLSPNSQLGSIPLSDGLSLTLQSLDLSNNSFNGLPVSFF 109

QY 119 ---QIRKSLTTDSQS---LGMGSTNMLLHVSSRRRLDEEPTISGTNNSVSGSNV 171  
Db 110 NARELFLDLSNMIIGETPSAIGDLHNL-----TLNLSMDALA-----GKLPTNLA 157

QY 172 SGNDNTVSGNHNHVSNNVTGSDNTVSGNHHVSGTKHIVTDNNNNVSG----- 224  
Db 158 SLRLNTVSLNNYFSGE-----IPGGRVV-----BFLDLSNLSLNGSLPPDFG 202

QY 225 --NDNNVSGSFHTVSGE-----HNTVSGSNNTVSG 252  
A;Residues: 1-773 <ROU>

Db 203 GYSQLYLVNSFNOISGEIPPEIGVNFPRNVTVDLSEFNLTG 243

RESULT 13  
T05322

hypothetical protein F18P4.240 - Arabidopsis thaliana  
N;Alternate names: hypothetical protein F1C12.60  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: T05322; T04898  
R;Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De  
ewes, H.W.; Mayer, K.F.X.; Schueller, C.  
submitted to the Protein Sequence Database, April 1998

A;Reference number: Z15408  
A;Accession: T05322  
A;Molecule type: DNA  
A;Residues: 1-1232 <BEV>  
A;Cross-references: UNIPROT:Q9SN91; EMBL:AL022224  
A;Experimental source: cultivar Columbia; BAC clone F1C12  
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, February 1998

A;Reference number: Z15388  
A;Accession: T04898  
A;Molecule type: DNA  
A;Residues: 1-305 <BEW>  
A;Cross-references: EMBL:AL021637  
A;Experimental source: cultivar Columbia; BAC clone F18P4  
C;Genetics:  
A;Map position: 4  
A;Introns: 863/1; 1116/1  
A;Note: F1C12.60; F18P4.240  
C;Superfamily: protein kinase Xa21; leucine-rich alpha-2-glycoprotein repeat homology; p

Query Match 11.7%; Score 162.5; DB 2; Length 1232;  
Best Local Similarity 25.3%; Pred. No. 0.00045;  
Matches 71; Conservative 41; Mismatches 94; Indels 75; Gaps 11;

QY 11 LNVFLGFILOVAGATSWCHDDHAL-----RGLAENLSKGAVRLRAAWS--GASCCSW 64  
Db 4 LVLLLLFILCFSGLGQPGIINNDLQTLLEVKKSLVTNPQDDPLR---QWNSDNINYCSW 60

QY 65 EGVGCE-TASGRVVALRPRKGLGGIIPSSIGELDLRLYLDLSCNSLVGEVPKSLQIRLK 123  
Db 61 TGVTCDNTGLFRVIALNLTLGLTGSISPMFCRFDNLHLDLSSNLLVGPITALS-NLT 119

QY 124 SLTTDSQSLGMSINNLLHVSSRRRLDEEPTISGTNN--SVGSGSNVNVSGNDNTVSG 181  
Db 120 SLES-----LFLFSNLTGEIPSQLGSLVNRSLRIGDNELVGDIPETL--G 164

QY 182 N-----NNHVSG-----SNNTVVTGSD 198  
Db 165 NLVNLQMLALASCRLTGPIPSQLGRVRYOSLIQDNYLEGPPIPAELGNCSDLTVFTAAE 224

QY 199 NTVGSGNHVSGTKHIVTDNNNNVSGNDNNVSGSFHTVSGE 239  
Db 225 NMLNGTIPAEIGR-----LENLEILNANSLTGEIPSQLGE 261

RESULT 14  
T00502

probable receptor-like protein kinase At2g23300 [imported] - Arabidopsis thaliana  
N;Alternate names: protein kinase homolog T20D16.7  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 16-Aug-2004  
C;Accession: T00502; A04623  
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul,  
submitted to the EMBL Data Library, November 1997

A;Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.  
A;Reference number: Z14164  
A;Accession: T00502  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-773 <ROU>



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2005, 20:15:44 ; Search time 43 Seconds  
(without alignments)  
463.519 Million cell updates/sec

Title: US-10-657-852A-15  
Perfect score: 1385  
Sequence: 1 MPEYMAKCMMLVFLGLFILO.....NTVSGNHNHIVSGSNKVVTDG 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCITUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfileesi.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 182   | 13.1        | 941    | 3     | US-09-336-447A-9   |
| 2          | 182   | 13.1        | 941    | 4     | US-09-952-267B-9   |
| 3          | 177   | 12.8        | 892    | 3     | US-09-336-447A-5   |
| 4          | 177   | 12.8        | 892    | 4     | US-09-952-267B-5   |
| 5          | 176.5 | 12.7        | 804    | 4     | US-10-101-464A-890 |
| 6          | 175   | 12.6        | 889    | 3     | US-09-336-447A-15  |
| 7          | 175   | 12.6        | 889    | 4     | US-09-952-267B-15  |
| 8          | 174   | 12.6        | 843    | 4     | US-10-101-464A-893 |
| 9          | 171.5 | 12.4        | 323    | 4     | US-10-101-464A-764 |
| 10         | 170.5 | 12.3        | 386    | 4     | US-10-101-464A-902 |
| 11         | 170.5 | 12.3        | 707    | 3     | US-09-228-986-80   |
| 12         | 170.5 | 12.3        | 707    | 4     | US-10-101-464A-80  |
| 13         | 170   | 12.3        | 624    | 3     | US-09-336-447A-7   |
| 14         | 170   | 12.3        | 624    | 4     | US-09-952-267B-7   |
| 15         | 168.5 | 12.2        | 187    | 4     | US-10-101-464A-732 |
| 16         | 164   | 11.8        | 828    | 4     | US-10-101-464A-934 |
| 17         | 163   | 11.8        | 692    | 4     | US-10-101-464A-897 |
| 18         | 158   | 11.4        | 942    | 4     | US-10-101-464A-911 |
| 19         | 155   | 11.2        | 867    | 4     | US-09-540-236-2676 |
| 20         | 153   | 11.0        | 653    | 4     | US-10-101-464A-953 |
| 21         | 152   | 11.0        | 224    | 4     | US-10-101-464A-770 |
| 22         | 152   | 11.0        | 257    | 4     | US-10-101-464A-743 |
| 23         | 150.5 | 10.9        | 854    | 4     | US-10-101-464A-896 |
| 24         | 150   | 10.8        | 666    | 3     | US-09-228-986-68   |
| 25         | 150   | 10.8        | 666    | 4     | US-10-101-464A-68  |
| 26         | 150   | 10.8        | 998    | 4     | US-10-101-464A-914 |
| 27         | 149.5 | 10.8        | 645    | 4     | US-09-345-473E-15  |

|    |       |      |     |   |                    |                   |
|----|-------|------|-----|---|--------------------|-------------------|
| 28 | 149.5 | 10.8 | 645 | 4 | US-09-345-473E-16  | Sequence 16, Appl |
| 29 | 149.5 | 10.8 | 873 | 3 | US-09-336-447A-13  | Sequence 13, Appl |
| 30 | 149.5 | 10.8 | 873 | 4 | US-09-952-267B-13  | Sequence 13, Appl |
| 31 | 149   | 10.8 | 878 | 4 | US-09-540-236-3401 | Sequence 3401, Ap |
| 32 | 147.5 | 10.6 | 711 | 3 | US-09-228-986-79   | Sequence 79, Appl |
| 33 | 147.5 | 10.6 | 711 | 4 | US-10-101-464A-79  | Sequence 79, Appl |
| 34 | 147.5 | 10.6 | 910 | 3 | US-09-228-986-72   | Sequence 72, Appl |
| 35 | 147.5 | 10.6 | 910 | 4 | US-10-101-464A-72  | Sequence 72, Appl |
| 36 | 146.5 | 10.6 | 635 | 4 | US-10-101-464A-932 | Sequence 932, App |
| 37 | 146   | 10.5 | 831 | 3 | US-09-336-447A-1   | Sequence 1, Appli |
| 38 | 146   | 10.5 | 831 | 4 | US-09-952-267B-1   | Sequence 1, Appli |
| 39 | 145.5 | 10.5 | 160 | 4 | US-10-101-464A-772 | Sequence 772, App |
| 40 | 145.5 | 10.5 | 370 | 4 | US-10-101-464A-944 | Sequence 944, App |
| 41 | 143.5 | 10.4 | 659 | 3 | US-09-228-986-75   | Sequence 75, Appl |
| 42 | 143.5 | 10.4 | 659 | 4 | US-10-101-464A-75  | Sequence 75, Appl |
| 43 | 143   | 10.3 | 144 | 4 | US-10-101-464A-739 | Sequence 739, App |
| 44 | 142.5 | 10.3 | 655 | 3 | US-09-228-986-70   | Sequence 70, Appl |
| 45 | 142.5 | 10.3 | 655 | 4 | US-10-101-464A-70  | Sequence 70, Appl |

## ALIGNMENTS

RESULT 1  
US-09-336-447A-9  
; Sequence 9, Application US/09336447A  
; Patent No. 6310190  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, ERIC J.  
; APPLICANT: AEBI, CHRISTOPH  
; APPLICANT: COPE, LESLIE D.  
; APPLICANT: MACIVER, ISOBEL  
; APPLICANT: FISKE, MICHAEL J.  
; APPLICANT: FREDENBURG, ROSS A.  
; TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS  
; FILE REFERENCE: AMCY:024  
; CURRENT APPLICATION NUMBER: US/09/336,447A  
; CURRENT FILING DATE: 1999-06-21  
; NUMBER OF SEQ ID NOS: 98  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 941  
; TYPE: PRT  
; ORGANISM: Moraxella catarrhalis  
US-09-336-447A-9

Query Match 13.1%; Score 182; DB 3; Length 941;  
Best Local Similarity 34.2%; Pred. No. 1e-08;  
Matches 39; Conservative 19; Mismatches 50; Indels 6; Gaps 1;  
QY 154 NTISGTNNVSGSGNNVSGNDNTVVGNNHVVSGNNTVVTGSDNTVVGNNHVVSGTKH 213  
Db 123 NEAMGEYTVAGANNQAKGNTVVGNGNKAIGNNTVVGNNQAKGEHSTIAGGK- 181  
QY 214 IVTDNNVVGNDNNVVGSHFTVSGHNTVVGNNHVVSGNNTVVGNNHVVSGKVVTDG 267  
Db 182 -----NNQATGNGSPAAGVENKADANNAVALGNKNTIEGTNSVAIGSNNTVTKG 230

RESULT 2  
US-09-952-267B-9  
; Sequence 9, Application US/09952267B  
; Patent No. 6753417  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, ERIC J.  
; APPLICANT: AEBI, CHRISTOPH  
; APPLICANT: COPE, LESLIE D.  
; APPLICANT: MACIVER, ISOBEL  
; APPLICANT: FISKE, MICHAEL J.  
; APPLICANT: FREDENBURG, ROSS A.  
; TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS  
; FILE REFERENCE: AMCY:024  
; CURRENT APPLICATION NUMBER: US/09/952,267B

; CURRENT FILING DATE: 2001-09-12  
 ; PRIOR APPLICATION NUMBER: US/09/336,447  
 ; PRIOR FILING DATE: 1999-06-21  
 ; NUMBER OF SEQ ID NOS: 98  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 9  
 ; LENGTH: 941  
 ; TYPE: PRT  
 ; ORGANISM: Moraxella catarrhalis  
 US-09-952-267B-9

Query Match 13.1%; Score 182; DB 4; Length 941;  
 Best Local Similarity 34.2%; Pred. No. 1e-08;  
 Matches 39; Conservative 19; Mismatches 50; Indels 6; Gaps 1;  
 QY 154 NTISGTVSGSGNNVSGNDNTVSGNNHVGSSNNTVTGSDNTVVGNNHVSGTKH 213  
 DB 123 NEAMGEYTVAGGANNQAKGNYSTVGGGNGKNAIGNNSTVVGSSNNQAKGEHSTIAGGK- 181  
 QY 214 IVTDNNVVGSDNNVSGSFHTVSGEHTVSGNNHVGSSNNTVSGNNHVGSSNKKVVDG 267  
 DB 182 -----NQATGGSFAGVENKADANNAVALGNKNTIEGTSVAIGSNNTVKTG 230

## RESULT 3

US-09-336-447A-5  
 ; Sequence 5, Application US/09336447A  
 ; Patent No. 6310190  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HANSEN, ERIC J.  
 ; APPLICANT: AEBI, CHRISTOPH  
 ; APPLICANT: COPE, LESLIE D.  
 ; APPLICANT: MACIVER, ISOBEL  
 ; APPLICANT: FISKE, MICHAEL J.  
 ; APPLICANT: FREDENBURG, ROSS A.  
 ; TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS  
 ; FILE REFERENCE: AMCY:024  
 ; CURRENT APPLICATION NUMBER: US/09/336,447A  
 ; CURRENT FILING DATE: 1999-06-21  
 ; NUMBER OF SEQ ID NOS: 98  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 892  
 ; TYPE: PRT  
 ; ORGANISM: Moraxella catarrhalis  
 US-09-336-447A-5

Query Match 12.8%; Score 177; DB 3; Length 892;  
 Best Local Similarity 31.8%; Pred. No. 2.9e-08;  
 Matches 35; Conservative 20; Mismatches 49; Indels 6; Gaps 1;  
 QY 158 GTNSVSGSGNNVSGNDNTVSGNNHVGSSNNTVTGSDNTVVGNNHVSGTKHIVTD 217  
 DB 143 GDSSTIGGGYNNQATGEKSTVAGGRNQATGNNSTVAGGSYNQATGNNSTVAGGSH---- 198  
 QY 218 NNNVSGNDNNVSGSFHTVSGEHTVSGNNHVGSSNNTVSGNNHVGSSNKKVVDG 267  
 DB 199 --NQATGGSFAGVENKANNNAVALGNKNTIDGNSVAIGSNNTIDSG 246

## RESULT 4

US-09-952-267B-5  
 ; Sequence 5, Application US/09952267B  
 ; Patent No. 6753417  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HANSEN, ERIC J.  
 ; APPLICANT: AEBI, CHRISTOPH  
 ; APPLICANT: COPE, LESLIE D.  
 ; APPLICANT: MACIVER, ISOBEL  
 ; APPLICANT: FISKE, MICHAEL J.  
 ; APPLICANT: FREDENBURG, ROSS A.  
 ; TITLE OF INVENTION: USPA1 AND USPA2 ANTIGENS OF MORAXELLA CATARRHALIS  
 ; FILE REFERENCE: AMCY:024

; CURRENT APPLICATION NUMBER: US/09/952,267B  
 ; CURRENT FILING DATE: 2001-09-12  
 ; PRIOR APPLICATION NUMBER: US/09/336,447  
 ; PRIOR FILING DATE: 1999-06-21  
 ; NUMBER OF SEQ ID NOS: 98  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 892  
 ; TYPE: PRT  
 ; ORGANISM: Moraxella catarrhalis  
 US-09-952-267B-5

Query Match 12.8%; Score 177; DB 4; Length 892;  
 Best Local Similarity 31.8%; Pred. No. 2.9e-08;  
 Matches 35; Conservative 20; Mismatches 49; Indels 6; Gaps 1;  
 QY 158 GTNSVSGSGNNVSGNDNTVSGNNHVGSSNNTVTGSDNTVVGNNHVSGTKHIVTD 217  
 DB 143 GDSSTIGGGYNNQATGEKSTVAGGRNQATGNNSTVAGGSYNQATGNNSTVAGGSH---- 198  
 QY 218 NNNVSGNDNNVSGSFHTVSGEHTVSGNNHVGSSNNTVSGNNHVGSSNKKVVDG 267  
 DB 199 --NQATGGSFAGVENKANNNAVALGNKNTIDGNSVAIGSNNTIDSG 246

## RESULT 5

US-10-101-464A-890  
 ; Sequence 890, Application US/10101464A  
 ; Patent No. 6768041  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Strabala, Timothy  
 ; APPLICANT: Nieuwenhuizen, Nicolaas  
 ; APPLICANT: Higgins, Colleen M.  
 ; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
 ; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
 ; FILE REFERENCE: 11000.1020C2  
 ; CURRENT APPLICATION NUMBER: US/10/101,464A  
 ; CURRENT FILING DATE: 2002-03-18  
 ; PRIOR APPLICATION NUMBER: 09/704,302  
 ; PRIOR FILING DATE: 2000-11-01  
 ; PRIOR APPLICATION NUMBER: 09/228,986  
 ; PRIOR FILING DATE: 1999-01-12  
 ; PRIOR APPLICATION NUMBER: 60/162,866  
 ; PRIOR FILING DATE: 1999-11-01  
 ; PRIOR APPLICATION NUMBER: PCT/US00/00724  
 ; PRIOR FILING DATE: 2000-01-11  
 ; NUMBER OF SEQ ID NOS: 989  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 890  
 ; LENGTH: 804  
 ; TYPE: PRT  
 ; ORGANISM: Pinus radiata  
 US-10-101-464A-890

Query Match 12.7%; Score 176.5; DB 4; Length 804;  
 Best Local Similarity 23.4%; Pred. No. 2.8e-08;  
 Matches 77; Conservative 54; Mismatches 105; Indels 93; Gaps 13;  
 QY 10 MLLVFLGFI-----QVAGATSWSCHD--DILHRLGLAE--NL 44  
 DB 6 LLLVMGVAMPHTHSQHTGGFTSVQRPFFNGRSMGKPSIAGYHEKRDVEALLSPKGIITL 65  
 QY 45 SGKGAVALRAAWGASCCSWEGVCETASGRVVALRPLKRGIGIIFSSIGELDHLAYLD 104  
 DB 66 DPYGLSNWTANNHNVCLWNGISCSPTNRVVEISLRYGLNGLTSPYIGNLSILRLHD 125  
 QY 105 LSGNSLVGEVPKSL-----QIRLSLTDSOS-----LGMGSINMLLHVSSRRTLDEPN 154  
 DB 126 LSSNALSGRIPAEFGQLKALRIILDSNNALTGSIPTCIGNGGLGTL-----DLDLN 179  
 QY 155 TISG-----TNNVSGSGNNVSGNDNTVSGNN--NHVSGSNNTV- 194  
 DB 180 AFSGRIPKELFNCTRLQRLIDLSHNSL-TGSIPTSIGNCALLQTLNTLNTFNLSGSIPTSLA 238

|                       |        |  |       |                                      |
|-----------------------|--------|--|-------|--------------------------------------|
| Query Match           | 12.6%  | Score 174;   | DB 4; | Length 843;                          |
| Best Local Similarity | 22.7%; | Pred. No. 5.2e-08;   |       |                                      |
| Matches               | 80;    | Conservative   | 55;   | Mismatches 107; Indels 110; Gaps 13; |
| Qy                    | 13     | VFLGFTLQVAGATSWSC-----BHDDLHALRGLAENLSGK--GAVRLRAAASGASCC      | 62    |                                      |
| Db                    | 15     | VMMGVAMPTDGTFSVESVPENGRRHKRDVEALLSPKESIIDPYGSLTNWTANNSHVC      | 74    |                                      |
| Qy                    | 63     | SWGVCETASGRVVALRLPKRGIGGIIPSSIGELDHLRYLDLSCNSLVEVPKSLQIRL      | 122   |                                      |
| Db                    | 75     | LWNGISCRPNTRKRWISLSEPCWNGTLSPYIGNLSLLRHLDSWNALSGRIPAEFG-QL     | 133   |                                      |
| Qy                    | 123    | KSLLTDSQSLSGMCGIN-----MLLHVSSRRLDEEPNTISGT--NNSVSGSGS-         | 167   |                                      |
| Db                    | 134    | KALRIIDLASGHLHGCIYPKELFNCTRLQRI DLSHNSFTGSIPTSIGNSALLQTMNLAE   | 193   |                                      |
| Qy                    | 168    | -----NNVVSQ-----NDNTVVSG-----                                  | 181   |                                      |
| Db                    | 194    | NQURGSIPAEFGRVLVHLESQLYNTLSLGSIPTSILANCTSLIKLELSDNN-LSGFIPESEF | 252   |                                      |
| Qy                    | 182    | -----NNNHVSGSNNTVVGT-----SDNTVVGSNNHVSGTKHIVTDNNNV             | 221   |                                      |

[illegible]

```

RESULT 9
US-10-101-464A-764
; Sequence 764, Application US/10101454A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 764
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-764

```

```

Query Match      12.4%; Score 171.5; DB 4; Length 323;
Best Local Similarity 25.5%; Pred. No. 2.4e-08;
Matches 82; Conservative 52; Mismatches 112; Indels 75; Gaps 14;

QY      10  MLIVFLGFIL-----QVAGATWSCHHD--DLHALRGIAE--NL  44
DB      6  LLIVMGVAMPTTSQHTGGTTSVQRPFPNGRSMWGKFSIAGYHEKRVDEALLSPFRKGITL  65
QY      45  SGKGAVRLRAAWGASCCSSEGVGCEATSGKRVVALRPLKRGLGIIIPSSIGELDHLRYLD  104
DB      66  DPYGLSNWTANNSHNVCLWNGISCSPTNRVVEISLRYGRLNGTLSPYIGNLSLLRHL  125
QY      105  LGSNSLVGEVPKSL---QIRLKSLLTDSQS-----LGMGSINMLLHVSRRTLDEPN  154
DB      126  LSSNALSGRIPAEFGQLKALRIILDSNALTGSIPTCIGNGLNGTL-----DLDSLWN  179
QY      155  TISG-----TNVSGSGSNVVSGNDNTVSGNN--NHVSGSNNTVV--  194
DB      180  AFSGRIPKELFNCTRQRIIDLNSNL-TGSIPT*SIGNCALLQTLNTGNYLSGSIPTSLA  238
QY      195  --TGSNDNTVVGSNHVSGTKHIVTDNNVVS-----GNDNNVSGSFHTVSGEHNT---VS  244
DB      239  NCTSLTDLVLSSNN-LSGP--IPSEFGLVSLKFLFDDNSIGSIPTSILVNCTSLFALK.  295
QY      245  GSNNTVVGSGNHIVSGSNKVVT  265
DB      296  GSGNKLSGPIPSVNLGMSIS  316

```

RESULT 10  
US-10-101-464A-902  
; Sequence 902, Application US/10101464A  
; Patent No. 6768041  
; GENERAL INFORMATION:  
; APPLICANT: Strbala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas

```

1  / APPLICANT: Higgins, Colleen M.
2  / TITLE OF INVENTION: Compositions Isolated from Plant Cells
3  / TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
4  / FILE REFERENCE: 11000.1020c2
5  / CURRENT APPLICATION NUMBER: US/10/101.464A
6  / CURRENT FILING DATE: 2002-03-18
7  / PRIOR APPLICATION NUMBER: 09/704,302
8  / PRIOR FILING DATE: 2000-11-01
9  / PRIOR APPLICATION NUMBER: 09/228,986
10 / PRIOR FILING DATE: 1999-01-12
11 / PRIOR APPLICATION NUMBER: 60/162,866
12 / PRIOR FILING DATE: 1999-11-01
13 / PRIOR APPLICATION NUMBER: PCT/US00/00724
14 / PRIOR FILING DATE: 2000-01-11
15 / NUMBER OF SEQ ID NOS: 989
16 / SOFTWARE: Fast-SEQ for Windows Version 4.0
17 / SEQ ID NO 902
18 / LENGTH: 386
19 / TYPE: PRT
20 / ORGANISM: Pinus radiata
21 / US-10-101-464A-902

```

|                       |                  |   |            |             |
|-----------------------|------------------|---|------------|-------------|
| Query Match           | 12.3%;           | Score 170.5;  | DB 4;      | Length 386; |
| Best Local Similarity | 32.3%;           | Pred. No. 3.8e-08;  |            |             |
| Matches               | 52; Conservative | 25; Mismatches  | 61; Indels | 23; Gaps 7; |
| Qy                    | 29               | CHHDDLHALRGLAENLSGKCAVPLRAAWS-GASCCSWEGVCETASGR----         | VVALRLPKR  | 84          |
| Dd                    | 38               | CNOSDVOALISFKOTVSSSSPLN-----WEVNRSCTWEGVTGGLSLQQFVKTLRLPGR  | 93         |             |
| Qy                    | 85               | GLGGIIPSSIGELDLHLRLDLSGNLVGEVPKSLQLRLKSLTTDSOSLGMGSINMLIHV- | 143        |             |
| Dd                    | 94               | RLRGITSDSLGRAHLREINVSNFLTIGEVPGLFT-LQHLL--EVLDSLFWNLSTVAPVA | 150        |             |
| Qy                    | 144              | ---SSRRTLDDEPNITSGTNNSVGSC-----SNNVWSG                      | 173        |             |
| Dd                    | 151              | OGLNSIRTFNISPRGVNPOLGSAVNLTFSFNVSNNSTFG                     | 191        |             |

RESULT 11

```

US-09-228-986-80
; Sequence 80, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions and
; TITLE OF INVENTION: and Their Use in
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228-986-80
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 1.0
; SEQ ID NO 80
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-80

```

|                       |        |  |                |                      |
|-----------------------|--------|--|----------------|----------------------|
| Query Match           | 12.3%; | Score 170.5;   | DB 3;          | Length 707;          |
| Best Local Similarity | 26.1%; | Pred. No. 8.9e-08;   |                |                      |
| Matches               | 81;    | Conservative 39;   | Mismatches 89; | Indels 101; Gaps 16; |
| QY                    | 9      | CMLLVFLGFILOVA---GATSWSCHHDDHLALRGLAENLSGKGAVALRAAWS--GASCC- | 62             |                      |
|                       |        | :   :  | :   :   :   :  |                      |
| Db                    | 8      | CLRLIITAIITASIATSHGTD---PDDVSALKGIYSSLNSPQOL---              | SCWSANGGDPCG   | 60                   |
| QY                    | 63     | -SWEGVCGCTAGSRVVALPLPRKGIGIIIPSSIGELDLHLYLDLGSNLVGVEPKSLQIR  | 121            |                      |
|                       |        | :   :   :   :  | :   :   :   :  |                      |
| Db                    | 61     | QSKWGVSCSGSS--VTLIKSLGLGSLYYQLSDLSLTLLDLSNNICGNIPYALPQK      | 118            |                      |
| QY                    | 122    | LKSLTTDSOSLGWGI-----NM-----LIH-----VSSRRTLDEBPNTI            | 156            |                      |
|                       |        | :   :   :   :  | :   :   :   :  |                      |

```

119 Db      LOELNLASNGLS-GTIPYSISNMTCGLTDLKLSHNLQSGOIQDIFQGLSLSLTDLDSFNTL 177
157 Qy      SGT-----NNSVSGSGNNVSVSGNDNTVSGNNHVS-----188
178 Db      TONLPQSPFSSLSLVLYLQNNQL-AGSVNVLANPLPLTDLNLTENNRFSGWVPNAWRSNQ 236
189 Qy      ---SNNVTVTG-----SDNTVSGSHVV---SCTKHIVTDNNNVSGNDNN 228
237 Db      FKYSGNSFATGPAPPPPPPTPPPPSNRPKSSNVVPSGSK-----CGNSNK 286
229 Qy      VSGSFHTVS 238
287 Db      KSLSGGAIVG 296

RESULT 12
US-10-101-464A-80
; Sequence 80, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-80

```

```

RESULT 13
US-09-336-447A-7
; Sequence 7, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP1 AND USP2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/336.447A
; PRIOR FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-336-447A-7

Query Match 12.3%; Score 170; DB 3; Length 624;
Best Local Similarity 30.6%; Pred. No. 8.3e-08;
Matches 38; Conservative 24; Mismatches 50; Indels 12; Gaps 2

Qy 154 NTISGTNNVSGSNVNVSGNDNTVTVSGNNHVSNNNTVTVGSDNTVTVGSNNHVSCTKH 213
Db 104 NQAKGEHSITAGGSNOATGRNSTVAGGSNNQAVGTNSTVAGGSNNQAKGANSFAAGVGN 163
Qy 214 IVTDNNHVSNDNNVSGSFHTVSGEHTVVS-----GNNTVSGSNHIVSG---SN 261
Db 164 QANTDNAVLGKNNTINGNNSAALGSENTVNNQKTVFILGNTNTNAQSGVLLGHETSG 223
Qy 262 KVVT 265
Db 224 KEAT 227

RESULT 14
US-09-952-267B-7
; Sequence 7, Application US/09952267B
; Patent No. 6753417
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP1 AND USP2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/952.267B
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US/09/336.447
; PRIOR FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-952-267B-7

```

|    |                          |               |                   |              |                 |             |
|----|--------------------------|---------------|-------------------|--------------|-----------------|-------------|
|    | Query Match              | 12.3%;        | Score 170;        | DB 4;        | Length 624;     |             |
|    | Best Local Similarity    | 30.6%;        | Pred.No. 8.3e-08; |              |                 |             |
|    | Matches 38; Conservative | 24;           | Mismatches 50;    | Indels 12;   | Gaps 2;         |             |
| Qy | 154                      | NTISGTTNNSVGS | GNNYVSGNDITTV     | SGNNHNVSGNN  | TVTGSDNTVVGSHNV | VSTGKH 213  |
|    |                          | : :           | :     :           | :     :      | :     :         | :           |
| Dp | 104                      | NOAKEHSHTLAGE | SNAQTGRNSTVAG     | NNQAQVGTNSTV | AGSSNNKAKGANSF  | PAAGVGN 163 |
|    |                          | :             | :                 | :            | :               | :           |



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2005, 20:21:56 ; Search time 138 Seconds  
(without alignments)  
645.441 Million cell updates/sec

Title: US-10-657-852a-15  
Perfect score: 1385  
Sequence: 1 MPEYMAKCMMLVFLGLQ.....NTVSGNHVSGSKVVDG 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1428581 seqs, 333598853 residues

Total number of hits satisfying chosen parameters: 1428581

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID                      | Description       |
|------------|-------|-------------|--------|-------------------------|-------------------|
| 1          | 262   | 18.9        | 1010   | 16 US-10-437-963-180288 | Sequence 180288,  |
| 2          | 206.5 | 14.9        | 1039   | 16 US-10-437-963-166082 | Sequence 166082,  |
| 3          | 204   | 14.7        | 1021   | 15 US-10-443-101-2      | Sequence 2, Appli |
| 4          | 198   | 14.3        | 284    | 16 US-10-437-963-148813 | Sequence 148813,  |
| 5          | 197   | 14.0        | 961    | 16 US-10-437-963-172660 | Sequence 172660,  |
| 6          | 194   | 14.0        | 917    | 16 US-10-437-963-115627 | Sequence 115627,  |
| 7          | 194   | 14.0        | 1164   | 16 US-10-437-963-192784 | Sequence 192784,  |
| 8          | 192.5 | 13.9        | 838    | 16 US-10-437-963-169754 | Sequence 169754,  |
| 9          | 192   | 13.9        | 1060   | 16 US-10-437-963-153194 | Sequence 153194,  |
| 10         | 190   | 13.7        | 1044   | 16 US-10-437-963-162505 | Sequence 162505,  |
| 11         | 188.5 | 13.6        | 1011   | 16 US-10-437-963-172329 | Sequence 172329,  |
| 12         | 188.5 | 13.6        | 1078   | 16 US-10-437-963-123119 | Sequence 123119,  |
| 13         | 186   | 13.4        | 278    | 15 US-10-424-599-267766 | Sequence 267766,  |

|    |       |      |      |    |                      |                   |
|----|-------|------|------|----|----------------------|-------------------|
| 14 | 186   | 13.4 | 676  | 16 | US-10-437-963-153306 | Sequence 153306,  |
| 15 | 185   | 13.4 | 1012 | 16 | US-10-437-963-192786 | Sequence 192786,  |
| 16 | 183.5 | 13.2 | 380  | 16 | US-10-437-963-136619 | Sequence 136619,  |
| 17 | 183.5 | 13.2 | 1133 | 16 | US-10-437-963-150876 | Sequence 150876,  |
| 18 | 183.5 | 13.2 | 1140 | 16 | US-10-437-963-119802 | Sequence 119802,  |
| 19 | 182   | 13.1 | 941  | 10 | US-09-952-267-9      | Sequence 9, Appli |
| 20 | 182   | 13.1 | 1231 | 16 | US-10-437-963-102913 | Sequence 102913,  |
| 21 | 181   | 13.1 | 653  | 15 | US-10-389-566-1625   | Sequence 1625, Ap |
| 22 | 181   | 13.1 | 1065 | 16 | US-10-437-963-109994 | Sequence 109994,  |
| 23 | 181   | 13.1 | 1469 | 16 | US-10-437-963-196963 | Sequence 196963,  |
| 24 | 180.5 | 13.0 | 675  | 16 | US-10-437-963-174437 | Sequence 174437,  |
| 25 | 180   | 13.0 | 1036 | 16 | US-10-437-963-149679 | Sequence 149679,  |
| 26 | 180   | 13.0 | 1080 | 16 | US-10-437-963-110291 | Sequence 110291,  |
| 27 | 179   | 12.9 | 516  | 15 | US-10-424-599-201138 | Sequence 201138,  |
| 28 | 178   | 12.9 | 153  | 15 | US-10-425-114-63986  | Sequence 63986, A |
| 29 | 178   | 12.9 | 792  | 16 | US-10-437-963-109998 | Sequence 109998,  |
| 30 | 178   | 12.9 | 1051 | 16 | US-10-437-963-156545 | Sequence 156545,  |
| 31 | 177.5 | 12.8 | 470  | 16 | US-10-437-963-172330 | Sequence 172330,  |
| 32 | 177.5 | 12.8 | 990  | 16 | US-10-437-963-170216 | Sequence 170216,  |
| 33 | 177   | 12.8 | 892  | 10 | US-09-952-267-5      | Sequence 5, Appli |
| 34 | 176.5 | 12.7 | 804  | 14 | US-10-101-464A-890   | Sequence 890, App |
| 35 | 176.5 | 12.7 | 804  | 17 | US-10-864-252-890    | Sequence 890, App |
| 36 | 176.5 | 12.7 | 817  | 16 | US-10-437-963-114827 | Sequence 114827,  |
| 37 | 176   | 12.7 | 852  | 15 | US-10-282-122A-62892 | Sequence 62892, A |
| 38 | 175.5 | 12.7 | 1036 | 16 | US-10-437-963-157688 | Sequence 157688,  |
| 39 | 175   | 12.6 | 889  | 10 | US-09-952-267-15     | Sequence 15, Appl |
| 40 | 175   | 12.6 | 1000 | 16 | US-10-437-963-138726 | Sequence 138726,  |
| 41 | 174.5 | 12.6 | 556  | 16 | US-10-437-963-109997 | Sequence 109997,  |
| 42 | 174.5 | 12.6 | 967  | 16 | US-10-437-963-118888 | Sequence 118888,  |
| 43 | 174.5 | 12.6 | 1060 | 16 | US-10-437-963-167059 | Sequence 167059,  |
| 44 | 174   | 12.6 | 344  | 16 | US-10-437-963-158951 | Sequence 158951,  |
| 45 | 174   | 12.6 | 843  | 14 | US-10-101-464A-893   | Sequence 893, App |

ALIGNMENTS

RESULT 1  
US-10-437-963-180288  
; Sequence 180288, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 180288  
; LENGTH: 1010  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_7766C.1.pep  
US-10-437-963-180288

Query Match 18.9%; Score 262; DB 16; Length 1010;  
Best Local Similarity 40.1%; Pred. No. 18-13;  
Matches 71; Conservative 19; Mismatches 53; Indels 34; Gaps 5;  
QY 7 KCML-LVFLGLQVAGATSWSCCHDDHLALRGLAENLS-GKGAVLRRAAAGSCCSW 64  
DB 8 RCLFSLVAFALLPPPPAAAAPCPEDLLALRAFAGNLSAGCGGAGLRAAAGDACCW 67  
QY 65 EGVGCTASGRVVALRPRKGLGIIPISSIGELDLRLYLDSGLVGEVPSLQIRLS 124

Db 68 DGVACD-RAARVTALRPGLEGIPPSLAALRQDLDSHNTLT-----113  
QY 125 LTTDSQSLGMSINMLLVHSSRRTLDEEPTISGT-----NNVSGSGNNVVSG 173  
Db 114 -----GGISALLAAVSLRTANLSSNLTLLDLAALPHLSAFNAGNSLSG 160

RESULT 2

US-10-437-963-166082  
; Sequence 166082, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 166082  
; LENGTH: 1039  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_64826C.1.pap  
US-10-437-963-166082

Query Match 14.9%; Score 206.5; DB 16; Length 1039;  
Best Local Similarity 29.1%; Pred. No. 6.6e-09;  
Matches 83; Conservative 52; Mismatches 103; Indels 47; Gaps 13;  
QY 5 MAKCCMLVFGFIQV--AGATSWCHDDHDLHALRGLAENLSGKGVRLRAAWSGAS-C 61  
Db 1 MARAMALLPCSYALALYSAGSSSSNATADLALLSPKSMSSP-SLGLMASWNSSHF 59  
QY 62 CSWEGVCETAS-GRVVALRPLKRGGLGIIIPSSIGELDLHRLYDLSGNSLVGVPKSI-- 118  
Db 60 CSWTGVCSSROQPEKVIQVIALQWNSCGLSRIISFFLGNLSFLTLGLGNQLVQGPSELGH 119  
QY 119 --QIRLKSLLTDSQSLGMSIN-----MLLVHSSRRTLDEEPTI-SGTNNVSGSG 166  
Db 120 LSKRLMLNST---NLLRGSIPVEMRGCTKMTLHLGNQLQGEIPAEIGSSKLNLINLY 176  
QY 167 -SNNVSGN-----DNTVSGNNHVSNNVTVTGSDN--TVVGSNNHVSNG--T 211  
Db 177 LTRNLLSGEIPQSLAEPLSLESLHSHNKLGEVPSALSNTLTNLNTRFNSNNMLSGVIPS 236  
QY 212 KHVITDNNVSGNDNNVSG-----SFFHTVSGEHNVTSGS 246  
Db 237 SLGMLPNLYELSLGFNNLSGPIPTSIWNISLRALSVOGNMLSGT 281

RESULT 3

US-10-443-101-2  
; Sequence 2, Application US/10443101  
; Publication No. US20040096941A1  
; GENERAL INFORMATION:  
; APPLICANT: MATSUBAYASHI, YOSHIKATSU  
; APPLICANT: SAKAGAMI, YUJI  
; TITLE OF INVENTION: RECEPTOR FOR PLANT CELL GROWTH FACTOR  
; FILE REFERENCE: 238013USO  
; CURRENT APPLICATION NUMBER: US/10/443,101  
; CURRENT FILING DATE: 2003-05-22  
; PRIOR APPLICATION NUMBER: JP 2002-335572  
; PRIOR FILING DATE: 2002-11-19

; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 2  
; LENGTH: 1021  
; TYPE: PRT  
; ORGANISM: Daucus carota  
US-10-443-101-2

Query Match 14.7%; Score 204; DB 15; Length 1021;  
Best Local Similarity 24.6%; Pred. No. 1.1e-08;  
Matches 81; Conservative 56; Mismatches 94; Indels 98; Gaps 14;

QY 10 MLLVFLGFIQV--VAGATSWCHDDHDLHAL-----RGLAENLSGKGVRLRAAWS----- 57  
Db 8 VILVGFQVQIVVNSQNTCSNDLKALEGFMRGLESSIDG-----WKWNESSS 58  
QY 58 -GASCCSWEGVCETA-----SGRVVALRPLKRGGLGIIIPSSIGELDLHRLYDL 106  
Db 59 FSSNCCDWGIGISCKSVSLGLDDVNESGRVVELELGRRLSGKLSVAKLDQLKVLNLT 118  
QY 107 GNSLVGVPKSL-----QIRLKSLLTDSQSLGMSINMLLVHSSRRTLDEEPTISG-- 158  
Db 119 HNSLGSIAAALLNLSLEVLDSLSSNDFS---GLFPLSLNLPRLVNLVNVENSFHLIPA 175  
QY 159 --TNN-----SVGSGSNNVSGNDNTV--VSGNNHVSNG-----SNNTVV 194  
Db 176 SLGNLPRIRIDLAMNYFDGSPVIGICNCSSVEYGLASNNLSGSIPOELFQLSNLSVL 235  
QY 195 TGSNTVTVGSNNHVSNGTKHIVTDNNVSGNDNNVSGSFHTV-----SGEHNVT 244  
Db 236 ALQNNRLSGA---LSSKLGKLSNLGRDLISSNKFCKIPDVFLELNKLMWYFSAQSNLFN 291  
QY 245 GS-----NNTVSGSNHI 256  
Db 292 GEMPRSLNSRSISLLSLRNNLTLSGGIYL 320

RESULT 4

US-10-437-963-148813  
; Sequence 148813, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 148813  
; LENGTH: 284  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_49206C.1.pap  
US-10-437-963-148813

Query Match 14.3%; Score 198; DB 16; Length 264;  
Best Local Similarity 27.6%; Pred. No. 6.2e-09;  
Matches 70; Conservative 42; Mismatches 116; Indels 26; Gaps 8;  
QY 5 MAKCCMLVFGFIQVAGATSWCHDDHDLHALRGLAENLSGKGVRLRAAW-SGASCCS 63  
Db 1 MQELQLLLVLAALVLLSFUSPATSCTEQEKSLQLFLRELSPSSSKFSRWSQSGTSCCT 60  
QY 64 WEGVGCETASGRVVALRPLKRGGLGIIIPSSIGELDLHRLYDLSGNSLVGVPKSLQIRLK 123

Db 61 WEIAC-GSNGVTLSLPSMALEGPIVSIAWLTGLRLDLSYNSLSGELPEL---IS 116  
QY 124 SLTDSQSLGMSINMLLVSSRRTIDEEPNTISGTNNVSGG---SNNVSGNDNTVVS 180  
Db 117 SASVAFDVSFNLRELQESS-PSLPHPLOVLNISHNFFPAGEFPTTWKKSDLVAIN 175  
QY 181 GNNHVSGSNTVVTGSDNTVGSNVHVGSKHIVTD-NNNVSGNDNNVSGSFHTVSGE 239  
Db 176 ASHTPFGA-----LPSSFCISSPSPAVLDLSYNLFGSIPAEIGKCSL--- 220  
QY 240 HNTVSGSNNTVSGS 253  
Db 221 -RVLKASNNNEINGS 233

## RESULT 5

US-10-437-963-172660  
; Sequence 172660, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 172660  
; LENGTH: 961  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_70775C.1.pap  
US-10-437-963-172660

Query Match 14.2%; Score 197; DB 16; Length 961;  
Best Local Similarity 29.5%; Pred. No. 4e-08;  
Matches 82; Conservative 102; Mismatches 102; Indels 56; Gaps 15;

QY 10 MLLVFLGFILOVAGATWS-CHDDHLALRGLAENLSGKAVRLRAWSGAS--CCSWEG 66  
Db 4 MOWLLLFMLLSRLSFSQTPNQDYSALQALMKWQNE-----PQSMWGSTDPCTSDG 58  
QY 67 VGCETASGRVVALRLPKRGLGGIIPSSIGELDHLRYLDLSGN-SLVGEVPKSLQIRLKS 125  
Db 59 ISC-SNGRVTEMRLSGINLQGLTSLNAIDQLSSLTLDLSNNLNGPDPPT-VNLKQL 115  
QY 126 TT-----DSQSLGMSINMLLVSSRRT-----LDEEPNTISG 158  
Db 116 TTLILGCSFTGDIPEQIGALRQLTFLALNSKNFTGGIPPTGLLSKLFWLDLSDNQLSG 175  
QY 159 TNNVSGSNNVSGNDNTVGSNNHVSNNVTVGTSDNTVGSNNHVSNGTKHIVTDN 218  
Db 176 -KIPVSSGSN---PGLDQLV---NAEHFFHSENLGTGPIDEKLFSEK---MNLHIVFDN 225  
QY 219 NNV---VSGNDNNVSGSFHTVSGEHTVSGSNNTVSGS 253  
Db 226 NNTGPIGSLGRVS-SIQIRLDHMQFSQ---PVPGS 259

## RESULT 6

US-10-437-963-115627  
; Sequence 115627, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 115627  
; LENGTH: 917  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_19205C.1.pap  
US-10-437-963-115627

Query Match 14.0%; Score 194; DB 16; Length 917;  
Best Local Similarity 26.3%; Pred. No. 6.8e-08;  
Matches 81; Conservative 43; Mismatches 108; Indels 76; Gaps 12;

QY 6 AKCCMLLVFLGFILOVAGATWSCHDDHLALRGLAENLSGKAVRLRAWS--GASCCS 63  
Db 4 ATAAULLLTAAILAAAGAV-----NDQVLAIVVFKSGVSDPGV--LAAMSEDAADRA 56  
QY 64 WGVGCGETASGRVVALRLPKRGLGGIIP-SSIGELDHLRYLDLSGNLSVGEVPKSLQIRL 122  
Db 57 WPGVSCDARAGRVDVALPFSAGLSGRLPRSAALLRLDALLSLALPGNNLSGPLPDALPPRA 116  
QY 123 KSLTTDSQSL-----GMGSINM-----LLHYSSRRTLDEEPNTISG 158  
Db 117 RALDISANSLSGYLPAALASCSSLSNLSGNLSPGVPDGIWSLPSLSKSLDLSGNQLAG 176  
QY 159 TNNVSGG-----SNNVSGNDNTVVS-----GNNHVSNNTV--V 194  
Db 177 ---SVPGGFPSSSLRVLDLSRLNLEGEIPADVGEAGLLKSLDVGHNLFTGELPESL 233  
QY 195 TGSDNTVVGSNHV-----VSGTKHIVTDNNNVSGNDNNVSGSFHTVSGEHTV 244  
Db 234 TGLSSIGAGGNALAGELPGWIGEMAALETLDLSGNRFVGAIPDGISGCKNLVE-----VD 288  
QY 245 GSNNTVSG 252  
Db 289 LSGNALTG 296

## RESULT 7

US-10-437-963-192784  
; Sequence 192784, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 192784  
; LENGTH: 1164  
; TYPE: PRT

|    |     |   |     |
|----|-----|---|-----|
| Qy | 127 | TDSQLGMSINMLLHVSSRRTLDEEPTISGTNNVSGSNN-----VWGSNDNTVVS      | 180 |
|    |     | :     :      :  |     |
| Db | 346 | -----IPSSGIW-----LDVSNRLGGELQEVDSSSDWPLOVLNISNLTGF          | 391 |
|    |     | :     :      :  |     |
| Qy | 181 | GNNHVGSSNNTVTGSDNTVVG--SNHVVGSTKHIVTD-----NNNVVG-----       | 224 |
|    |     | :     :      :  |     |
| Db | 392 | PSTTWKMSNLVAINGASNSTGHI PSSFCISLSFAALDLCYNQFSGEI PAGIGKCSAL | 451 |
|    |     | :     :      :  |     |
| Qy | 225 | -----NDNNVSGS-----FHTVSGEHTVSGSNNTVSGS-NHIVSGSKKVTD         | 266 |
|    |     | :     :      :  |     |
| Db | 452 | RMLKAGHNINSGALPDLFIATSLSEY--LGFPNNGQGITKLVIKLNLFVD          | 502 |
|    |     | :     :      :  |     |

```

RESULT 9
US-10-437-963-153194
; Sequence 153194, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)/B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 153194
; LENGTH: 1060
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_53170C.1.pep
US-10-437-963-153194

```

```

; LENGTH: 1060
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_53170C.1.pep
US-10-437-963-153194

Query Match      13.9%   Score 192;   DB 16;   Length 1060;
Best Local Similarity 27.5%; Pred. No. 1.2e-07;
Matches 72; Conservative 44; Mismatches 112; Indels 34; Gaps 9;

QY          11 LLVLFGFTLVAGATSWCHHDDLHALRGLAENL-----SGKGAVRLRAAWGASCC 62
              :|||:::|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db          12 VIVLLLLLVLKNGAAQRCSGGDLAALRGFSAGLDGVGDGWPAVGNASSSTSDGGDCC 71
              :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY          63 SWEGVCGETASGRVVALRLPKRGLGGTIPSSIGELPHLRDYLDLSGNSLNVEVPKSLQIRL 122
              :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db          72 AWRGVACDEA-GEVGVGVLPNATLRGVVAESLAGLAALVNLSSNALRGALPAGI-LRL 129
              :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY          123 KSLTTDSQL-----GMGSINMLAHVSRRRTLDEEPTNTISGTNNVSGSNNVVSGNDNTV 178
              :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db          130 RALQVLVDVSVNALEGAVAAAAYVDLPAMREFNVSYNAFNG-SHPVLAGAGRUTSYD---- 184
              :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY          179 VSGNN--NHVSGSNNTVTGSDNTVCGSNHVSVG-----TKGHVTDNNNVSGN 225
              :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db          185 VSGNSFAGHDAAALCGASPGLRTRLNSMNGSFDPFPVFGQCRSLVELSDIGNAIAGAL 244
              :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY          226 DNNVSG--SPHTVSGEHNVTSG 245
              :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db          245 PDDVFGLTSLQVLSLHTNSLSG 266
              :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

```

RESULT 10  
US-10-437-963-162505  
; Sequence 162505, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 162505  
LENGTH: 1044  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_61590C.1.p  
US-10-437-963-162505

Query Match 13.7%; Score 190; DB 16; Length 1044;  
Best Local Similarity 26.5%; Pred. No. 1.8e-07;  
Matches 79; Conservative 47; Mismatches 114; Indels 58; Gaps 14;  
QY 1 MPEYMAKCCMLLVFLG-FILOVAGATSWCHHD-DLHALRGLAENLSG-KGAVRLRAAWS 57  
DB 1 MIRLFASCPKLIPLAVFIFSSSLPLAISDDTDTREALLCFKQISDPNGAL---SSWT 57  
QY 58 GASC--CSWEGVGCETASG--RVVALRLPKRGLGGIIPSSIGELDLHRLYDLDSNLSVGE 113  
DB 58 NTSLNFCWQGVSCNTPQQLRWALNVSKGUGLIPPCIGNLSSIASLDJLNNAPLKG 117  
QY 114 VPKSL---QIRLKSITTDQSGLMGSINNMLL-HVSSRRRLDEPNTISGTNNNSVSG-- 166  
DB 118 IPSELGHLQISVNL-----SINSLGHIPELDSCLSKVLVSLCNNSLOGEIP 167  
QY 167 -----SNNVSGNDNT-----VVSNNHNVSGNNTVVTGSDNTV-- 201  
DB 168 PSITQCTHLQVVLCCNKLQGRITKFGMLHEKLTLDLSNALTGDIPPLGSSPSFVTV 227  
QY 202 -VGSNNHVSCTKHIVTDNN--VVSNDNNVSGSFHTV---SGEHTVSGNNTVSGS 253  
DB 228 DLGGNQLTGGIPEFLANSSSQLVLSLQNNLTGIPPSIFNSSKLTIYLRNRLVGS 285

RESULT 11  
US-10-437-963-172329  
; Sequence 172329, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 172329  
; LENGTH: 1011  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_70476C.1.p  
US-10-437-963-172329

Query Match 13.6%; Score 188.5; DB 16; Length 1011;  
Best Local Similarity 30.5%; Pred. No. 2.3e-07;

Matches 74; Conservative 32; Mismatches 88; Indels 49; Gaps 12;  
QY 11 LLVFLGFILQVAGATSWCHHDLDHALRGLAENLSGKGA VR-----LRAAWSGAS----- 60  
DB 4 MSLSLTHALLLTQVTSQTINGDDLSAL-----LSFKSLIRDDPREVMSSWDTAGNTNM 57  
QY 61 -----CCSWEGVGCETA--SGRVVALRLPKRGLGGIIPSSIGELDLHRLYDLDSNLSVGEV 114  
DB 58 PAPVICQMTGVSNNRRHPRVTTTLRLSGAGLVGTISPOLGNLTLHRLVLDLSANSLDGI 117  
QY 115 PKSL---QIRLKSITTDQSGLMGSINNMLLHVSSRRRLDEPNTISGTNNNSVSGSNV 170  
DB 118 PASLGGCKRLTYLNLSNHL---GSPDDLGSSKLAIFD-----VGHNNLTGVNPKSP 169  
QY 171 VSGNDNTVVSNNHNVSGNNTVVTGSDNTVVSNNHVSCTKHIVTDNNNNVSGNDNNVS 230  
DB 170 --SNLTTLV---KPIIETN--PIDGKDLNMGNN---LTSLTHFVLEGNRFT---GNIP 214  
QY 231 GSF 233  
DB 215 ESF 217  
RESULT 12  
US-10-437-963-123119  
; Sequence 123119, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 123119  
; LENGTH: 1078  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1078)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_25984C.1.p  
US-10-437-963-123119

Query Match 13.6%; Score 188.5; DB 16; Length 1078;  
Best Local Similarity 28.5%; Pred. No. 2.5e-07;  
Matches 76; Conservative 32; Mismatches 92; Indels 67; Gaps 12;  
QY 12 LVFLGFILQVAGAT-----SWSCHDDHLHALRGLAENLSGKGA VRRAAWSGA-SCCSWEG 66  
DB 8 VVSGVLVLVLAATICGRAACAERAAALGFLAEASPAGDAIVADWRGSPDCCRWDG 67  
QY 67 VGCETA---SGRVVALRLPKRGLGGIIPSSIGELDLHRLYDLDSNLSVGEVVK----- 116  
DB 68 VCGGAGDGDGAVTRLSLGRGFNGTISFIGNLTGLTHLNSGNLAGGFPEVLSLPN 127  
QY 117 -----SLQIRLKSITTDQSGLMGSINNMLLHVSSRRRLDEPNTI-----S 157  
DB 128 VTVDVSVNCLSGELPSVATGAAGGLSLE-VLDVSSNLLAGQFPFSAIWEHTPRLVSLN 186  
QY 158 GTNNNSVSG-----SNNVSGNDNTVVS---GNNHVV-----SGSNNTVV 194  
DB 187 ASNNSPHSGIPSLCVSCPALAVLDLSNVNLSG----VISPGFNCQSLRVLNLRNN--L 240



```
Db 122 -----GFAIE-VNVSSNGFTGPHAFPGAPNLTVDITGNAFSGGINVTA 167
QY 178 -----VMSGNNHVSQS-----NNTVVVTS----- 197
Db 168 LCASPVKVLRFSANAFSGDVPAGFGOCKLLNDLFLDGNGLTGS LPKDYMPALRKL SIQ 227
QY 198 DNTVGSNNHVSQTKHIVTD---NNNVSGNDNNVSG-----SFHTVSGEHN----- 241
Db 228 ENKLSGLDDDLGNLTETITQIDLSYNNFNGNIPDVFGLRSLRSLNLSNQLNGTLPLSL 287
QY 242 -----TVSGSNNTVSGS-----NHIVSGSNKV 263
Db 288 SSCFMLRVVSLRNSLSGEITIDCRLLTRLNPFDA GTNKL 327
```

Search completed: May 9, 2005, 20:34:42  
Job time : 144 secs

*This Page Blank (uspto)*

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2005, 20:10:27 ; Search time 73 Seconds  
(without alignments)  
1414.589 Million cell updates/sec

Title: US-10-657-852a-15  
Perfect score: 1385  
Sequence: 1 MPEYAKCCMLLVFLGLFIQ.....NTVSGSNHIVGSKNVKVTIDG 267

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 1385  | 100.0       | 267    | 8     | Adm41473 Perennial |
| 2          | 1290  | 93.1        | 269    | 8     | Adm41474 Tall fesc |
| 3          | 992.5 | 71.7        | 262    | 8     | Adm41481 Tall fesc |
| 4          | 980.5 | 70.8        | 254    | 8     | Adm41480 Perennial |
| 5          | 973.5 | 70.3        | 256    | 8     | Adm41482 Perennial |
| 6          | 910.5 | 65.7        | 277    | 8     | Adm41477 Perennial |
| 7          | 908.5 | 65.6        | 277    | 8     | Adm41476 Tall fesc |
| 8          | 904.5 | 65.3        | 281    | 8     | Adm41483 Perennial |
| 9          | 901.5 | 65.1        | 281    | 8     | Adm41475 Perennial |
| 10         | 901.5 | 65.1        | 285    | 8     | Adm41479 Tall fesc |
| 11         | 884   | 63.8        | 280    | 8     | Adm41478 Perennial |
| 12         | 668   | 48.2        | 243    | 8     | Adm41472 Tall fesc |
| 13         | 655   | 47.3        | 243    | 8     | Adm41471 Perennial |
| 14         | 530   | 38.3        | 118    | 2     | Aay22472 Grass ant |
| 15         | 258.5 | 18.7        | 1036   | 5     | Abb93804 Herbicide |
| 16         | 204   | 14.7        | 1021   | 8     | Adoi4158 Carrot ph |
| 17         | 197.5 | 14.3        | 598    | 3     | Aag20654 Arabidops |
| 18         | 197.5 | 14.3        | 599    | 3     | Aag20653 Arabidops |
| 19         | 197.5 | 14.3        | 751    | 5     | Abb93984 Herbicide |
| 20         | 190.5 | 13.8        | 1008   | 8     | Adn72587 Thale cre |
| 21         | 182   | 13.1        | 941    | 2     | Aaw68205 M. catarr |
| 22         | 181   | 13.1        | 653    | 8     | Adj49621 Oll-assoc |
| 23         | 177   | 12.8        | 892    | 2     | Aaw68203 M. catarr |
| 24         | 176.5 | 12.7        | 735    | 3     | Aag32217 Arabidops |
| 25         | 176.5 | 12.7        | 743    | 3     | Aag32216 Arabidops |

|    |       |      |      |   |          |                    |
|----|-------|------|------|---|----------|--------------------|
| 26 | 176.5 | 12.7 | 804  | 3 | AAB25515 | Aab25515 Pinus rad |
| 27 | 176   | 12.7 | 852  | 6 | ABU34968 | Abu34968 Protein e |
| 28 | 175   | 12.6 | 889  | 2 | AAW68208 | Aaw68208 M. catarr |
| 29 | 174   | 12.6 | 843  | 3 | AAB25518 | Aab25518 Pinus rad |
| 30 | 173.5 | 12.5 | 590  | 5 | AB91206  | Abb91206 Herbicide |
| 31 | 171.5 | 12.4 | 323  | 3 | AAB25445 | Aab25445 Pinus rad |
| 32 | 170.5 | 12.3 | 386  | 3 | AAB25527 | Aab25527 Pinus rad |
| 33 | 170.5 | 12.3 | 707  | 3 | AAB25112 | Aab25112 Pinus rad |
| 34 | 170   | 12.3 | 624  | 2 | AAW68204 | Aaw68204 M. catarr |
| 35 | 168.5 | 12.2 | 187  | 3 | AAB25413 | Aab25413 Pinus rad |
| 36 | 168.5 | 12.2 | 1095 | 5 | AB91562  | Abb91562 Herbicide |
| 37 | 166.5 | 12.0 | 938  | 8 | ADQ36979 | Adq36979 Cell prol |
| 38 | 166.5 | 12.0 | 938  | 8 | ADQ15649 | Adq15649 Rice stre |
| 39 | 166   | 12.0 | 966  | 5 | ABB93915 | Abb93915 Herbicide |
| 40 | 166   | 12.0 | 966  | 8 | ADI32622 | Adi32622 Thale cre |
| 41 | 165.5 | 11.9 | 854  | 4 | AAW42228 | Aam42228 Soybean r |
| 42 | 165.5 | 11.9 | 854  | 4 | AAW42230 | Aam42230 Soybean r |
| 43 | 165.5 | 11.9 | 854  | 4 | AAW42232 | Aam42232 Soybean r |
| 44 | 165   | 11.9 | 627  | 5 | ABB92188 | Abb92188 Herbicide |
| 45 | 165   | 11.9 | 627  | 8 | ADN72847 | Adn72847 Thale cre |

ALIGNMENTS

RESULT 1  
ADM41473  
ID ADM41473 standard; protein; 267 AA.  
XX ADM41473;  
AC ADM41473;

DT 03-JUN-2004 (first entry)  
XX Perennial ryegrass antifreeze protein AFP2.  
DE Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;  
XX Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;  
KW antitox; litholytic; nephrotropic; cytotostatic.  
XX Lolium perenne.

Key Location/Qualifiers  
FT Peptide 1..23  
FT Region 60..71  
FT Region /label= Signal\_peptide  
FT Region /note= "Conserved Cys pairs identified N-terminal of leucine-rich repeats of receptor-like kinases"  
FT Region 76..117  
FT Region /note= "Conserved leucine-rich repeats"  
FT Region 152..266  
FT Region /note= "7-amino acid sequence repeat region identified in antifreeze proteins"

WO2004022700-A2.

18-MAR-2004.

09-SEP-2003; 2003WO-NZ000199.

09-SEP-2002; 2002US-0409557P.

(GENE-) GENESIS RES & DEV CORP LTD.  
(WRIG-) WRIGHTSON SEEDS LTD.

Demmer D) Shenk WA, Hall C, Fish SA;  
WPI; 2004-248453/23.

N-PSDB; ADM41460.

XX New antifreeze proteins and encoding polynucleotides, useful for  
modulating cold tolerance in organisms, as food additives, or for  
treating tumors or disorders associated with the presence of unwanted  
biocrystals (e.g. gout).

PS Claim 7; SEQ ID NO 16; 71pp; English.

XX The present sequence is that of APP2, an antifreeze protein of perennial

CC ryegrass. The encoding cDNA was isolated from a leaf blade cDNA

CC expression library. The invention provides forage grass (perennial

CC ryegrass and tall fescue) antifreeze proteins and the polynucleotides

CC encoding them ADM41458-ADM41483. The polynucleotides were isolated from

CC tissues taken at different times of year (winter and spring) and from

CC different parts of the plants. The polynucleotides can be used to

CC modulate the cold tolerance of an organism, especially plants, mammals,

CC insects, fungi, archaea and bacteria. The method involves incorporating

CC an antifreeze polynucleotide, under the control of a gene promoter

CC sequence, into the genome of the organism, or introducing double-stranded

CC RNA corresponding to the polynucleotide into the cells of the organism,

CC thereby inhibiting expression of an antifreeze polypeptide. The

CC antifreeze protein can be used for the cryopreservation of a cell or

CC tissue, as a food additive of a frozen food product, in a method for

CC decreasing the time required to dehydrate a composition, to treat a

CC disorder characterised by biocrystals associated with disorders such as

CC gout and kidney stones, to preserve the viability of a molecular biology

CC reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and

CC to protect a plant from damage due to frost or freezing.

XX Sequence 267 AA;

Query Match 100.0%; Score 1385; DB 8; Length 267;

Best Local Similarity 100.0%; Pred. No. 5.4e-114;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPEYMAKCMMLVFLGFILOVAGATSWSCHDDDLHALRGLAENLSGKGAVALRAAWSCAS 60

DB 1 MPEYMAKCMMLVFLGFILOVAGATSWSCHDDDLHALRGLAENLSGKGAVALRAAWSCAS 60

QY 61 CCSWEGVGCETASGRVVALRPLPKRGLGIIIPSSIGELDHLYLDLSGNSLVGEVPSKLSQI 120

DB 61 CCSWEGVGCETASGRVVALRPLPKRGLGIIIPSSIGELDHLYLDLSGNSLVGEVPSKLSQI 120

QY 121 RLKSLTTDSQSLGMSINMLLVHVSRRRLDDEPNTISGTNNVSGSGNNVSGNDNTVVS 180

DB 121 RLKSLTTDSQSLGMSINMLLVHVSRRRLDDEPNTISGTNNVSGSGNNVSGNDNTVVS 180

QY 181 GNNNHVSGSNNVTVTGSDNTVVGSHVVGSGTKHIVTDNNNVSGNDNNVSGSFHTVSGEH 240

DB 181 GNNNHVSGSNNVTVTGSDNTVVGSHVVGSGTKHIVTDNNNVSGNDNNVSGSFHTVSGEH 240

QY 241 NTVSGSNNTVSGSNHIVSGSNKVVTDG 267

DB 241 NTVSGSNNTVSGSNHIVSGSNKVVTDG 267

RESULT 2

ADM41474

ID ADM41474 standard; protein; 269 AA.

XX AC ADM41474;

XX DT 03-JUN-2004 (first entry)

XX DE Tall fescue antifreeze protein.

XX KW Antifreeze; fescue; cold tolerance; transgenic; plant; antitout;

XX KW litholytic; nephrotropic; cytostatic.

XX OS Schedonorus arundinaceus.

XX FH Key Location/Qualifiers

FT Peptide 1..24

FT /label= signal\_peptide

FT Region 61..72

FT /note= "Conserved Cys pairs identified N-terminal of

FT leucine-rich repeats of receptor-like kinases"

FT Region 77..118

FT /note= "Conserved leucine-rich repeats"

FT Misc-difference 124

FT /note= "Encoded by TAG"

FT Region

FT 154..268

FT /note= "7-amino acid sequence repeat region identified in

FT antifreeze proteins"

PN WO2004022700-A2.

XX 18-MAR-2004.

XX 09-SEP-2003; 2003WO-NZ000199.

XX 09-SEP-2002; 2002US-0409557P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX (WRIG-) WRIGHTSON SEEDS LTD.

XX Demmer J, Shenk MA, Hall C, Fish SA;

XX WPI; 2004-248453/23.

XX N-PSDB; ADM41461.

XX New antifreeze proteins and encoding polynucleotides, useful for

XX modulating cold tolerance in organisms, as food additives, or for

XX treating tumors or disorders associated with the presence of unwanted

XX biocrystals (e.g. gout).

XX Claim 7; SEQ ID NO 17; 71pp; English.

CC The present sequence is that of an antifreeze protein of tall fescue. The

CC encoding cDNA was isolated from a leaf blade cDNA expression library. The

CC invention provides forage grass (perennial ryegrass and tall fescue)

CC antifreeze proteins and the polynucleotides encoding them ADM41458-

CC ADM41483. The polynucleotides were isolated from tissues taken at

CC different times of year (winter and spring) and from different parts of

CC the plants. The polynucleotides can be used to modulate the cold

CC tolerance of an organism, especially plants, mammals, insects, fungi,

CC archaea and bacteria. The method involves incorporating an antifreeze

CC polynucleotide, under the control of a gene promoter sequence, into the

CC genome of the organism, or introducing double-stranded RNA corresponding

CC to the polynucleotide into the cells of the organism, thereby inhibiting

CC expression of an antifreeze polypeptide. The antifreeze protein can be

CC used for the cryopreservation of a cell or tissue, as a food additive of

CC a frozen food product, in a method for decreasing the time required to

CC dehydrate a composition, to treat a disorder characterised by biocrystals

CC associated with disorders such as gout and kidney stones, to preserve the

CC viability of a molecular biology reagent, to destroy unwanted tissue in a

CC patient e.g. tumour tissue, and to protect a plant from damage due to

CC frost or freezing.

XX Sequence 269 AA;

Query Match 93.1%; Score 1290; DB 8; Length 269;

Best Local Similarity 93.3%; Pred. No. 1.4e-105;

Matches 251; Conservative 6; Mismatches 10; Indels 2; Gaps 2;

QY 1 MPEYMAKCMMLVFLGFI-LOVAGATSWSCHDDDLHALRGLAENLSGKGAVALRAAWSCA 59

DB 1 MPEYMAKCMMLLALLAFILLOVAGATSWSCHDDDLRALRGLFAENLSGKGAVALRAAWSCA 60

QY 60 SCCSWEGVGCETASGRVVALRPLPKRGLGIIIPSSIGELDHLYLDLSGNSLVGEVPSKLSQ 119

DB 61 SCCSWEGVGCETASGRVVALRPLPKRGLGIIIPSSIGELDHLYLDLSGNSLVGEVPSKLSQ 120

QY 120 IRLKSLTTDSQSLGMSINMLLVHVS-RRYLDDEPNTISGTNNVSGSGNNVSGNDNTV 178

DB 121 IRLKSLTTDSQSLGMSINMLLVHVS-RRYLDDEPNTISGTNNVSGSGNNVSGNDNTV 180

QY 179 VSGNNNHVSGSNNVTVTGSDNTVVGSHVVGSGTKHIVTDNNNVSGNDNNVSGSFHTVSG 238

DB 181 ISGNNNHVSGSNNVTVTGSDNTVVGSHVVGSGTKHIVTDNNNVSGNDNNVSGSFHTVSG 240

QY 239 EHNTVSGSNNTVSGSNHIVSGSNKVVTDG 267

|    |   |
|----|---|
| XX | Stability of a molecular biology reagent, to destroy unwanted tissue in a |
| PT | New antifreeze proteins and encoding polynucleotides, useful for          |



The present sequence is that of APP3, an antifreeze protein of perennial ryegrass. The invention provides forage grass (perennial ryegrass and tall fescue) antifreeze proteins and the polynucleotides encoding them (ADMA1458-ADMA1483). The polynucleotides were isolated from tissues taken at different times of year (winter and spring) and from different parts of the plants. The polynucleotides can be used to modulate the cold tolerance of an organism, especially plants, mammals, insects, fungi, archaea and bacteria. The method involves incorporating an antifreeze polynucleotide, under the control of a gene promoter sequence, into the genome of the organism, or introducing double-stranded RNA corresponding to the polynucleotide into the cells of the organism, thereby inhibiting expression of an antifreeze polypeptide. The antifreeze protein can be used for the cryopreservation of a cell or tissue, as a food additive of a frozen food product, in a method for decreasing the time required to dehydrate a composition, to treat a disorder characterised by bicrystals associated with disorders such as gout and kidney stones, to preserve the viability of a molecular biology reagent, to destroy unwanted tissue in a

DR N-PSDB; ADM41463.  
 XX  
 XX New antifreeze proteins and encoding polynucleotides, useful for  
 PT modulating cold tolerance in organisms, as food additives, or for  
 FT treating tumors or disorders associated with the presence of unwanted  
 PT biocrystals (e.g. gout).  
 XX  
 XX  
 PS Claim 7; SEQ ID NO 19; 71pp; English.  
 XX  
 XX The present sequence is that of an antifreeze protein of tall fescue. The  
 CC invention provides forage grass (perennial ryegrass and tall fescue).  
 CC antifreeze proteins and the polynucleotides encoding them ADM41458-  
 CC ADM41483. The polynucleotides were isolated from tissues taken at  
 CC different times of year (winter and spring) and from different parts of  
 CC the plants. The polynucleotides can be used to modulate the cold  
 CC tolerance of an organism, especially plants, mammals, insects, fungi,  
 CC archaea and bacteria. The method involves incorporating an antifreeze  
 CC polynucleotide, under the control of a gene promoter sequence, into the  
 CC genome of the organism, or introducing double-stranded RNA corresponding  
 CC to the polynucleotide into the cells of the organism, thereby inhibiting  
 CC expression of an antifreeze polypeptide. The antifreeze protein can be  
 CC used for the cryopreservation of a cell or tissue, as a food additive of  
 CC a frozen food product, in a method for decreasing the time required to  
 CC dehydrate a composition, to treat a disorder characterised by biocrystals  
 CC associated with disorders such as gout and kidney stones, to preserve the  
 CC viability of a molecular biology reagent, to destroy unwanted tissue in a  
 CC patient e.g. tumour tissue, and to protect a plant from damage due to  
 CC frost or freezing.  
 CC  
 SQ Sequence 277 AA;

Query Match 65.6%; Score 908.5; DB 8; Length 277;  
 Best Local Similarity 67.9%; Pred. No. 86-72; Mismatches 19; Gaps 4;  
 Matches 198; Conservative 19; Indels 19; Gaps 4;  
 QY 5 MAKCCMLLVFLGFIQVAGATSWCHDDHLAGRLAENLSCGKAVRLRAAWGASCCSW 64  
 DB 1 MAKCCMLLVFLVLLPAASAT--SCHPDDRLALRGFVGNLNG--GGVLLHGAWSGLCCAW 57  
 QY 65 EGVGCETASGRVVALRLP-----KRLGGIIPSSIGELDLHLRYLDLSGNS 109  
 DB 58 EGVGCDGTSGRVVALRLPISLDCGKLSLNLANERLVGTIPSWIGELDHCHVLDSDNS 117  
 QY 110 LVGEVPSLQIRKSLITDQSLGMSINMLLHV--SSRRTLDEPNTISGTNNVSGNSN 168  
 DB 118 LVGKVPNSLQIRKSLITDQSLGMSINMLLHV--SSRRTLDEPNTISGTNNVSGNSN 177  
 QY 169 NVVSGNDNTVSGNNHVSNNVSGNNVTVGSDNTVSGNNHVSNGTRHIVTDNNNVSGNDNN 228  
 DB 178 NAVSGNDNTVSGNNHVSNNVSGNNVTVGSDNTVSGNNHVSNGTRHIVTDNNNVSGNDNN 237  
 QY 229 VSGSPHTVSGHNTVSGNNHVSNNVSGNNVTVGSDNTVSGNNHVSNGTRHIVTDNNNVSGNDNN 265  
 DB 238 VSGSHTVSGHNTVSGNNHVSNNVSGNNVTVGSDNTVSGNNHVSNGTRHIVTDNNNVSGNDNN 274

RESULT 8  
 ADM41483  
 ID ADM41483 standard; protein; 281 AA.  
 XX  
 AC ADM41483;  
 XX  
 XX 03-JUN-2004 (first entry)  
 XX  
 XX Perennial ryegrass antifreeze protein.  
 XX  
 XX Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;  
 KW antigout; litholytic; nephrotropic; cytostatic.  
 XX  
 OS Lolium perenne.  
 XX  
 XX Key Location/Qualifiers  
 FH Peptide 1..19  
 FT

FT Region /label= Signal\_peptide  
 FT 10..22  
 FT /note= "Conserved lipoprotein membrane attachment site"  
 FT 53..64  
 FT /note= "Conserved Cys-pairs identified N-terminal or  
 FT leucine-rich repeats of receptor-like kinases"  
 FT 69..125  
 FT /note= "Conserved leucine-rich repeats"  
 FT 165..279  
 FT /note= "7-amino acid sequence repeat region identified in  
 FT antifreeze proteins"  
 WO2004022700-A2.  
 18-MAR-2004.  
 09-SEP-2003; 2003WO-NZ000199.  
 09-SEP-2002; 2002US-0409557P.  
 (GENE-) GENESIS RES & DEV CORP LTD.  
 (WRIG-) WRIGHTSON SEEDS LTD.  
 Demmer J, Shenk MA, Hall C, Fish SA;  
 WPI; 2004-248453/23.  
 N-PSDB; ADM41470.  
 New antifreeze proteins and encoding polynucleotides, useful for  
 modulating cold tolerance in organisms, as food additives, or for  
 treating tumors or disorders associated with the presence of unwanted  
 biocrystals (e.g. gout).  
 Claim 7; SEQ ID NO 26; 71pp; English.

The present sequence is that of an antifreeze protein of perennial  
 ryegrass. The invention provides forage grass (perennial ryegrass and  
 tall fescue) antifreeze proteins and the polynucleotides encoding them  
 ADM41458-ADM41483. The polynucleotides were isolated from tissues taken  
 at different times of year (winter and spring) and from different parts  
 of the plants. The polynucleotides can be used to modulate the cold  
 tolerance of an organism, especially plants, mammals, insects, fungi,  
 archaea and bacteria. The method involves incorporating an antifreeze  
 polynucleotide, under the control of a gene promoter sequence, into the  
 genome of the organism, or introducing double-stranded RNA corresponding  
 to the polynucleotide into the cells of the organism, thereby inhibiting  
 expression of an antifreeze polypeptide. The antifreeze protein can be  
 used for the cryopreservation of a cell or tissue, as a food additive of  
 a frozen food product, in a method for decreasing the time required to  
 dehydrate a composition, to treat a disorder characterised by biocrystals  
 associated with disorders such as gout and kidney stones, to preserve the  
 viability of a molecular biology reagent, to destroy unwanted tissue in a  
 patient e.g. tumour tissue, and to protect a plant from damage due to  
 frost or freezing.

Sequence 281 AA;

Query Match 65.3%; Score 904.5; DB 8; Length 281;  
 Best Local Similarity 66.5%; Pred. No. 1.8e-71;  
 Matches 187; Conservative 20; Mismatches 51; Indels 23; Gaps 5;  
 QY 5 MAKCCMLLVFLGFIQVAGATSWCHDDHLAGRLAENLSCGKAVRLRAAWGASCCSW 64  
 DB 1 MAKCCMLLVFLVLLPAASAT--SCHLDDRLALRGFVGNLNGAL--LRGTWSGSCCDW 57  
 QY 65 EGVGCETASGRVVALRLP-----KRLGGIIPSSIGELDLHLRYLDLSGNS 109  
 DB 58 EGVGCDGTSGRVVALRLPISLDCGKLSLNLANERLVGTIPSWIGELDHCHVLDSDNS 117  
 QY 110 LVGEVPSLQIRKSLITDQSLGMSINMLLHV--SSRRTLDEPNTISGTNNVSG 164  
 DB 118 LVGKVPNSLQIRKSLITDQSLGMSINMLLHV--SSRRTLDEPNTISGTNNVSG 177

|           |   |
|-----------|---|
| CC        | expression of an antifreeze polypeptide. The antifreeze protein can be used for the cryopreservation of a cell or tissue, as a food additive of a frozen food product, in a method for decreasing the time required to dehydrate a composition, to treat a disorder characterised by biocrystals associated with disorders such as gout and kidney stones, to preserve the viability of a molecular biology reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and to protect a plant from damage due to frost or freezing. |
| XX        |   |
| SQ        | Sequence 281 AA;  |
|           | Query Match<br>Best Local Similarity 65.1%; Score 901.5; DB 8; Length 281;<br>Matches 186; Conservative 21; Mismatches 51; Indels 23; Gaps 5;   |
| Qy        | 5 MAKCCMLLVFLGFIQVAGATSWSCHHDDLHALRGLAENLSKGAVRLRAAWSGASCSSW 64   |
| Db        | 1 MAKCWILLFLFVLLAMSAT--SCHLDDLRALRGFGVNLNGGAL-LRGTWSGSCCDW 57   |
| Qy        | 65 EGVGCFTASGRVVVALRP-----KRGIGGIIPSSIIGELDHLRYLDLSGNS 109  |
| Db        | 58 EGVGCDSGRVTALRPIPLEDCGKLKSLANLANERLVGTIPSWGELDHHCYLVLSDNS 117  |
| Qy        | 110 LVGEVP----KSLOIRLKSLTTDSQISLGMSINMLLHV-SSRETLDEEPTNISGTWNVSVC 164   |
| Db        | 118 LVGKAPNLSLNSLOIRLKGLATAGRSIGMAFANMPLHVGKNRRRTLDEQTNTIHGTNTVR 177  |
| Qy        | 165 SGNNNVSGNDNTVWGNNNHVSGSNNTVTVTGSNDNTVSGSNHHVVSCTKHIVTDNNNVSG 224  |
| Db        | 178 SGNDNAVSGNDNTVICGNNTVSGSNNTTAGSDGDNIVTSGSNHIVCGTKHIITDNNNDVSG 237   |
| Qy        | 225 NDNNVSGSFHTVSGEHNTVSGSNNTVSGSNHHIVSNGNKVVT 265  |
| Db        | 238 NDNNVSGSFHTVSGSHNTVSGSNNTVSGSNHHVVSNGNKLV 278   |
| RESULT 10 |   |
| ID        | ADM41479 standard; protein; 285 AA.   |
| XX        |   |
| AC        | ADM41479;   |
| XX        |   |
| DT        | 03-JUN-2004 (first entry)   |
| XX        |   |
| DB        | Tall fescue antifreeze protein.   |
| XX        |   |
| KW        | Antifreeze; fescue; cold tolerance; transgenic; plant; antigenic;   |
| KW        | litholytic; nephrotropic; cytosolic.  |
| XX        |   |
| OS        | Schedonorus arundinaceus.   |
| PH        | Location/Qualifiers   |
| FT        | Peptide 1..17   |
| FT        | /label= Signal_peptide  |
| FT        | Region 12..22   |
| FT        | /notes= "Conserved lipoprotein membrane attachment site"  |
| FT        | Region 52..63   |
| FT        | /note= "Conserved Cys-pairs identified N-terminal or leucine-rich repeats of receptor-like kinases"   |
| FT        | Region 68..134  |
| FT        | /note= "Conserved leucine-rich repeats"   |
| FT        | Domain 134..158   |
| FT        | /note = Transmembrane domain  |
| FT        | Region 170..284   |
| FT        | /note= "7-amino acid sequence repeat region identified in antifreeze proteins"  |
| XX        |   |
| FN        | WO2004022700-A2.  |
| XX        |   |
| PD        | 18-MAR-2004.  |
| XX        |   |
| PF        | 09-SEP-2003; 2003WO-NZ000199.   |
| XX        |   |

```

PR 09-SEP-2002; 2002US-0409557P.
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA (WRIG-) WRIGHTSON SEEDS LTD.
XX
XX Demmer J, Shenk MA, Hall C, Fish SA;
XX WPI; 2004-248453/23.
XX N-PSDB; ADM41466.
XX
XX New antifreeze proteins and encoding polynucleotides, useful for
XX modulating cold tolerance in organisms, as food additives, or for
XX treating tumors or disorders associated with the presence of unwanted
XX biocrystals (e.g. gout).
XX
XX Claim 7; SEQ ID NO 22; 71pp; English.
XX
XX The present sequence is that of an antifreeze protein of tall fescue. The
XX encoding cDNA was isolated from a leaf blade cDNA expression library. The
XX invention provides forage grass (perennial ryegrass and tall fescue)
XX antifreeze proteins and the polynucleotides encoding them ADM41458-
XX ADM41483. The polynucleotides were isolated from tissues taken at
XX different times of year (winter and spring) and from different parts of
XX the plants. The polynucleotides can be used to modulate the cold
XX tolerance of an organism, especially plants, mammals, insects, fungi,
XX archaea and bacteria. The method involves incorporating an antifreeze
XX polynucleotide, under the control of a gene promoter sequence, into the
XX genome of the organism, or introducing double-stranded RNA corresponding
XX to the polynucleotide into the cells of the organism, thereby inhibiting
XX expression of an antifreeze polypeptide. The antifreeze protein can be
XX used for the cryopreservation of a cell or tissue, as a food additive of
XX a frozen food product, in a method for decreasing the time required to
XX dehydrate a composition, to treat a disorder characterised by biocrystals
XX associated with disorders such as gout and kidney stones, to preserve the
XX viability of a molecular biology reagent, to destroy unwanted tissue in a
XX patient e.g. tumour tissue, and to protect a plant from damage due to
XX frost or freezing.
XX
XX Sequence 285 AA;
XX
XX Query Match 65.1%; Score 901.5; DB 8; Length 285;
XX Best Local Similarity 65.5%; Pred. No. 3.5e-71;
XX Matches 186; Conservative 17; Mismatches 54; Indels 27; Gaps 3;
XX
QY 10 MLLVFLGFLQVAGATSWCHDDHLALRLGLAENLSGKGAVALRAWGSACCSWEGVGC 69
Db :||:||||:| :||:||||:| :||:||||:| :||:||||:| :||:||||:|
QY 3 LLLLFLGLPACATSSCHPDDLRALRGFAKNVGG-GVLLRTAWSGTSCWEGVGC 61
Db :||:||||:| :||:||||:| :||:||||:| :||:||||:| :||:||||:|
QY 70 ETASGRVALRLPKRGLG-----GIIPSSIGELDHLRYLD 104
Db :||:||||:| :||:||||:| :||:||||:| :||:||||:| :||:||||:|
QY 62 NGASGRITTLPLPRGLAGTITGASLAGLARLESLNANRLVGTIPSGWIGELDHLRYLD 121
Db :||:||||:| :||:||||:| :||:||||:| :||:||||:| :||:||||:|
QY 105 LSGNSLVGEVPKSLQIRLKSITDQSQSGMSINMLHVS-SRTILDEEPTISGTNNVS 163
Db :||:||||:| :||:||||:| :||:||||:| :||:||||:| :||:||||:|
QY 122 LSHNSLVGEVPLNRLQIRLKLTTTGHLGMAFTNNPLDKNRRKTLAQPTISGTNNLV 181
Db :||:||||:| :||:||||:| :||:||||:| :||:||||:| :||:||||:|
QY 164 GSGNNVVGNDNTVSGNNHVSNNNTVTGSDNTVSGNNHVSNGTKHVTIDNNNVS 223
Db :||:||||:| :||:||||:| :||:||||:| :||:||||:| :||:||||:|
QY 182 LSGRNNVVGNDNTVISENNNTVSGNFVTITGSDNVLTVSGNNHVSNGSHVITDNNNVS 241
Db :||:||||:| :||:||||:| :||:||||:| :||:||||:| :||:||||:|
QY 224 GNDNNVSGFHTVSGEHTVSGNNNTVSGNNHVSNGSHVITDNNNVS 267
Db :||:||||:| :||:||||:| :||:||||:| :||:||||:| :||:||||:|
QY 242 GDDNNVSGFHKVSGSHNTVSGNNNTVSGNNHVSNGSHVITDNNNVS 285
Db :||:||||:| :||:||||:| :||:||||:| :||:||||:| :||:||||:|
XX
XX RESULT 11
XX ADM41478
XX ID ADM41478 standard; protein; 280 AA.
XX AC ADM41478;
XX XX
XX DT 03-JUN-2004 (first entry)
XX
XX

```

```

DE Perennial ryegrass antifreeze protein AFP4.
XX Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;
KW Antigout; litholytic; nephrotropic; cytostatic.
XX
XX Lolium perenne.
XX
XX Key Location/Qualifiers
XX Peptide 1..22
XX /label= Signal_peptide
XX Region 12..22
XX /note= "Conserved lipoprotein membrane attachment site"
XX Region 52..63
XX /note= "Conserved Cys-pairs identified N-terminal or
XX leucine-rich repeats of receptor-like kinases"
XX Region 68..134
XX /note= "Conserved leucine-rich repeats"
XX Domain 134..153
XX /note = Transmembrane domain
XX Region 165..279
XX /note= "7-amino acid sequence repeat region identified in
XX antifreeze proteins"
XX
XX WO2004022700-A2.
XX
XX 18-MAR-2004.
XX
XX 09-SEP-2003; 2003WO-NZ000199.
XX
XX 09-SEP-2002; 2002US-0409557P.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX (WRIG-) WRIGHTSON SEEDS LTD.
XX
XX Demmer J, Shenk MA, Hall C, Fish SA;
XX
XX WPI; 2004-248453/23.
XX N-PSDB; ADM41465.
XX
XX New antifreeze proteins and encoding polynucleotides, useful for
XX modulating cold tolerance in organisms, as food additives, or for
XX treating tumors or disorders associated with the presence of unwanted
XX biocrystals (e.g. gout).
XX
XX Claim 7; SEQ ID NO 21; 71pp; English.
XX
XX The present sequence is that of AFP4, an antifreeze protein of perennial
XX ryegrass. The encoding cDNA was isolated from a leaf and pseudostem cDNA
XX expression library. The invention provides forage grass (perennial
XX ryegrass and tall fescue) antifreeze proteins and the polynucleotides
XX encoding them ADM41458-ADM41483. The polynucleotides were isolated from
XX tissues taken at different times of year (winter and spring) and from
XX different parts of the plants. The polynucleotides can be used to
XX modulate the cold tolerance of an organism, especially plants, mammals,
XX insects, fungi, archaea and bacteria. The method involves incorporating
XX an antifreeze polynucleotide, under the control of a gene promoter
XX sequence, into the genome of the organism, or introducing double-stranded
XX RNA corresponding to the polynucleotide into the cells of the organism,
XX thereby inhibiting expression of an antifreeze polypeptide. The
XX antifreeze protein can be used for the cryopreservation of a cell or
XX tissue, as a food additive of a frozen food product, in a method for
XX decreasing the time required to dehydrate a composition, to treat a
XX disorder characterised by biocrystals associated with disorders such as
XX gout and kidney stones, to preserve the viability of a molecular biology
XX reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and
XX to protect a plant from damage due to frost or freezing.
XX
XX Sequence 280 AA;
XX
XX Query Match 63.8%; Score 884; DB 8; Length 280;
XX Best Local Similarity 65.5%; Pred. No. 1.2e-69;
XX Matches 186; Conservative 14; Mismatches 52; Indels 32; Gaps 4;
XX

```

```
QY 10 MLLVFLGFILOVAGATWSCHDDHLALRGIAENLSGKGAURLRAAWSGASCCSWGVGC 69
D 3 LLLFLFLAFLPVACAATSSCHPDDLALRGFAKNLGG-CGVLLRTAWSGTSCCVWEGVGC 61
QY 70 ETASGRVALRLPKRGIG-----GIIPSSIGELDLRLYLD 104
D 62 NGASGRVTLWLPRGLAGTITGASLAGLARLSLNNRLVGTIPSWIGELDLRLYLD 121
QY 105 LSGNSLVGEVPSKLIQRLKSLTDSOSLGMGSINMLLHVS-SRRTLDEEPNTISGTNNVS 163
D 122 LSHNSLVGELP-----NLKGLTTTGHLLGWAFSMPLDVKPNRTLAVQPNITISGTNNVS 176
QY 164 GSGSNVNVSGNDNTVSGNNHVSNNVSGNTVTGSDNTVSGSNHVSNGTGHIVTDNNVVS 223
D 177 LSGRNNVTSGNDNTVSGNNHVSNNVSGNTVTGSDNTVSGSNHVSNGRHHIVTDNNVVS 236
QY 224 GDNVNVSGSPHTVSGRHHNTVSGSNHVSNNVSGSNHVSNNKVVTDG 267
D 237 GDDNVNVSGSPHKVSGSNHVSNNVSGSNHVSNNKVVTDG 280

RESULT 12
ADM41472
ID ADM41472 standard; protein; 243 AA.
AC ADM41472;
XX
DT 03-JUN-2004 (first entry)
DE Tall fescue antifreeze protein.
XX
KW Antifreeze; fescue; cold tolerance; transgenic; plant; antitoxin;
KW litholytic; nephrotropic; cytosolic.
XX
OS Schedonorus arundinaceus.
XX
FH Key Location/Qualifiers
FT Peptide 1..20 /label= Signal_peptide
FT Region 53..64 /note= "Conserved Cys-pairs identified N-terminal or
FT /note= "leucine-rich repeats of receptor-like kinases"
FT Region 107..242 /note= "7-amino acid sequence repeat region identified in
FT antifreeze proteins"
XX
PN WO2004022700-A2.
XX
PD 18-MAR-2004.
XX
PF 09-SEP-2003; 2003WO-NZ000199.
XX
PR 09-SEP-2002; 2002US-0409557P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (WRIG-) WRIGHTSON SEEDS LTD.
XX
PI Demmer J, Shenk MA, Hall C, Fish SA;
XX
DR WPI; 2004-248453/23.
DR N-PSDB; ADM41459.
XX
PT New antifreeze proteins and encoding polynucleotides, useful for
PT modulating cold tolerance in organisms, as food additives, or for
PT treating tumors or disorders associated with the presence of unwanted
PT biocrystals (e.g. gout).
XX
PS Claim 7; SEQ ID NO 15; 71pp; English.
XX
CC The present sequence is that of an antifreeze protein of tall fescue. The
CC encoding cDNA was isolated from a basal stem cDNA expression library. The
CC invention provides forage grass (perennial ryegrass and tall fescue)
CC antifreeze proteins and the polynucleotides encoding them ADM41458-
```

```
CC ADM41483. The polynucleotides were isolated from tissues taken at
CC different times of year (winter and spring) and from different parts of
CC the plants. The polynucleotides can be used to modulate the cold
CC tolerance of an organism, especially plants, mammals, insects, fungi,
CC archaea and bacteria. The method involves incorporating an antifreeze
CC polynucleotide, under the control of a gene promoter sequence, into the
CC genome of the organism, or introducing double-stranded RNA corresponding
CC to the polynucleotide into the cells of the organism, thereby inhibiting
CC expression of an antifreeze polypeptide. The antifreeze protein can be
CC used for the cryopreservation of a cell or tissue, as a food additive of
CC a frozen food product, in a method for decreasing the time required to
CC denature a composition, to treat a disorder characterised by biocrystals
CC associated with disorders such as gout and kidney stones, to preserve the
CC viability of a molecular biology reagent, to destroy unwanted tissue in a
CC patient e.g. tumour tissue, and to protect a plant from damage due to
CC frost or freezing.
XX
SQ Sequence 243 AA;

Query Match 48.2%; Score 668; DB 8; Length 243;
Best Local Similarity 53.0%; Pred. No. 1.2e-50;
Matches 151; Conservative 17; Mismatches 53; Indels 64; Gaps 5;

QY 5 MAKCCMLVFLGFILOVAGATWSCHDDHLALRGIAENLSGKGAURLRAAWSGASCCSW 64
D 1 MAKCCMLVFLGFILOVAGATWSCHDDHLALRGIAENLSGKGAURLRAAWSGASCCSW 64
QY 65 EGVGCTASGRVALRLPKRGIGIIPSSIGELDLRLYLDLGSNLVGEVPSKLIQRLKLS 124
D 58 EGAGCDGASGRVT-----SFQILLKG 78

QY 125 LTTDSOSLGMGSINMLLHV-SRRTLDEEPNTISGTNNVSGSNHVSNNVSGNDNTVSGNN 183
D 79 LTTAGSLGKAFNPLHVKSQGLTDEEHTITGINNTVSGSNHVSNNVSGNDNTVSGNN 138
QY 184 NHVSGSNNTVTGSDNTVSGSNHVSNGTGHIVTDNNNVSGNDNTVSGSFHTVSGHH--- 240
D 139 NVVSGSHNTVFGDGNFLSGSNHVSNGRHHVTDNKNVSGDHTVSGSNTVSGNHHII 198
QY 241 -----NTVSGSNNTVSGSNHVSNNVSGSNHVSNNKVVTDG 267
D 199 SASHSTISGNNHNTVSGSNHVSNNVSGSNHVSNNKVVTDG 243

RESULT 13
ADM41471
ID ADM41471 standard; protein; 243 AA.
AC ADM41471;
XX
DT 03-JUN-2004 (first entry)
DE Perennial ryegrass antifreeze protein AFP1.
XX
KW Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;
KW antitoxin; litholytic; nephrotropic; cytosolic.
XX
OS Lolium perenne.
XX
FH Key Location/Qualifiers
FT Peptide 1..20 /label= Signal_peptide
FT Region 13..23 /note= "Conserved lipoprotein membrane attachment site"
FT Region 53..64 /note= "Conserved Cys-pairs identified N-terminal or
FT /note= "leucine-rich repeats of receptor-like kinases"
FT Region 107..242 /note= "7-amino acid sequence repeat region identified in
FT antifreeze proteins"
XX
PN WO2004022700-A2.
XX
```

PD 18-MAR-2004.  
 XX 09-SEP-2003; 2003WO-NZ000199.  
 PF XX 09-SEP-2002; 2002US-0409557P.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (WRIG-) WRIGHTSON SEEDS LTD.  
 XX Demmer J, Shenk MA, Hall C, Fish SA;  
 XX WPI; 2004-248453/23.  
 DR N-PSDB; ADM41458.  
 XX New antifreeze proteins and encoding polynucleotides, useful for  
 PT modulating cold tolerance in organisms, as food additives, or for  
 PT treating tumors or disorders associated with the presence of unwanted  
 PT biocrystals (e.g. gout).  
 XX Claim 7; SEQ ID NO 14; 71pp; English.  
 PS The present sequence is that of AFP1, an antifreeze protein of perennial  
 XX ryegrass. The encoding cDNA was isolated from a pseudostem cDNA  
 CC expression library. The invention provides forage grass (perennial  
 CC ryegrass and tall fescue) antifreeze proteins and the polynucleotides  
 CC encoding them ADM41458-ADM41483. The polynucleotides were isolated from  
 CC tissues taken at different times of year (winter and spring) and from  
 CC different parts of the plants. The polynucleotides can be used to  
 CC modulate the cold tolerance of an organism, especially plants, mammals,  
 CC insects, fungi, archaea and bacteria. The method involves incorporating  
 CC an antifreeze polynucleotide, under the control of a gene promoter  
 CC sequence, into the genome of the organism, or introducing double-stranded  
 CC RNA corresponding to the polynucleotide into the cells of the organism,  
 CC thereby inhibiting expression of an antifreeze polypeptide. The  
 CC antifreeze protein can be used for the cryopreservation of a cell or  
 CC tissue, as a food additive of a frozen food product, in a method for  
 CC decreasing the time required to dehydrate a composition, to treat a  
 CC disorder characterised by biocrystals associated with disorders such as  
 CC gout and kidney stones, to preserve the viability of a molecular biology  
 CC reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and  
 CC to protect a plant from damage due to frost or freezing.  
 XX Sequence 243 AA;  
 SQ  
 Query Match 47.3%; Score 655; DB 8; Length 243;  
 Best Local Similarity 51.9%; Pred. No. 1.8e-49;  
 Matches 148; Conservative 18; Mismatches 55; Indels 64; Gaps 5;  
 QY 5 MAKCCMLLVFGFILOVAGATSWCHHDDLHALRGLAENLSGKGAIVRLRAWGSACCSW 64  
 DB 1 MAKCCMLLVFGFILOVAGATSWCHHDDLHALRGLAENLSGKGAIVRLRAWGSACCSW 64  
 QY 65 EGVGCTASGRVVALRLPKRGLGGIIIPSSIGELDHLRYLDLSGNSLVEVPKSLQIRLKS 124  
 DB 58 EGVGCDGASGCV-----KSPQILLKG 78  
 QY 125 LTTDSQSLGMSGINMLLHVS-SRRTLDEPNTISGTNNVSGSGNNVSGNDNTVVSNN 183  
 DB 79 LTAAGRSILGKAFTEPHLVKESQGLDEHDNTIIGINNVSGSGNNVSGNDNTVVSNN 138  
 QY 184 NHVSGSNNTVVGSDNTVGSNNHVSQTKHIVTDNNNNVSGNDNNVSGSFHTVSGHNTV 243  
 DB 139 NVVSGSHNTVVGSDNTVGSNNHVSQTKHIVTDNNNNVSGNDNNVSGSFHTVSGHNTV 198  
 QY 244 SGS-----NNTVSGSNHIVSGSNKVVTDG 267  
 DB 199 SGSHSTVSGHNTVSGSNHVSQTKHIVTDNNNNVSGNDNNVSGSFHTVSGHNTV 243  
 RESULT 14  
 AAY22472  
 ID AAY22472 standard; protein; 118 AA.  
 XX

AC AAY22472;  
 XX 29-SEP-1999 (first entry)  
 DT XX Grass anti-freeze protein sequence.  
 DE XX Anti-freeze protein; grass; plant; frozen food product; frost tolerance;  
 KW frozen confectionery.  
 XX OS Lolium perenne.  
 XX PN WO9937782-A2.  
 XX PD 29-JUL-1999.  
 XX PF 23-DEC-1998; 98WO-EP008553.  
 XX PR 22-JAN-1998; 98GB-00001408.  
 XX PA (UNIL ) UNILEVER NV.  
 XX PI (UNIL ) UNILEVER PLC.  
 XX Jarman CD, Sidebottom CM, Twigg S, Worrall D;  
 WPI; 1999-458697/38.  
 DR N-PSDB; AAX99717.  
 XX New plant anti-freeze protein useful in frozen food products.  
 PT Claim 3; Page 36; 39pp; English.  
 PS This sequence is the plant anti-freeze protein of the invention. The anti-  
 CC -freeze protein is characterised in that at least 40% of its amino acids  
 CC are from the group of serine, threonine and asparagine. The anti-freeze  
 CC protein can be used in frozen food products, especially frozen  
 CC confectionery. Anti-freeze proteins are especially used in food products,  
 CC which are heated, e.g. by pasteurisation, blanching or sterilisation  
 CC prior to freezing. Plants transformed with a nucleic acid sequence  
 CC encoding the anti-freeze protein have an increased frost tolerance. Prior  
 CC art anti-freeze proteins have not been applied to commercially available  
 CC food products, due to high costs and complicated processes for obtaining  
 CC the protein. Also prior art anti-freeze proteins have tended to  
 CC destabilise during processing especially during the pasteurisation step.  
 CC This is overcome by the present anti-freeze protein. The anti-freeze  
 CC proteins provide an ice particle size following an ice recrystallisation  
 CC inhibition assay of 15 µm M or less. The anti-freeze protein ingredient  
 CC means that mixes can be frozen under quiescent conditions, e.g. in a shop  
 CC or home freezer without the formation of unacceptable ice crystal shapes  
 CC and hence with a texture different to products normally obtained via  
 CC quiescent freezing  
 XX Sequence 118 AA;  
 SQ  
 Query Match 38.3%; Score 530; DB 2; Length 118;  
 Best Local Similarity 85.5%; Pred. No. 7.3e-39;  
 Matches 100; Conservative 10; Mismatches 7; Indels 0; Gaps 0;  
 QY 150 DEPNITISGTNNVSGSGNNVSGNDNTVVSNNHVSQTKHIVTDNNNNVSGNDNNVSGSFHTVSGHNTVSGSNKVVTD 266  
 DB 1 DEPNITISGTNNVSGSGNNVSGNDNTVVSNNHVSQTKHIVTDNNNNVSGNDNNVSGSFHTVSGHNTVSGSNKVVTD 60  
 QY 210 GTHKIVTDNNNNVSGNDNNVSGSFHTVSGHNTVSGSNHVSQTKHIVTDNNNNVSGNDNNVSGSFHTVSGHNTVSGSNKVVTD 266  
 DB 61 GTHKIVTDNNNNVSGNDNNVSGSFHTVSGHNTVSGSNHVSQTKHIVTDNNNNVSGNDNNVSGSFHTVSGHNTVSGSNKVVTD 117  
 RESULT 15  
 ABB93804  
 ID ABB93804 standard; protein; 1036 AA.  
 XX  
 AC ABB93804;  
 XX 31-MAY-2002 (first entry)  
 DT

XX Herbicidally active polypeptide SEQ ID NO 3015.  
DE  
XX  
XX Herbicidal; plant; agriculture; herbicide.  
KW  
XX  
XX Arabidopsis thaliana.  
OS  
XX WO200210210-A2.  
PN  
XX  
XX 07-FEB-2002.  
PD  
XX  
XX 28-AUG-2001; 2001WO-EP009892.  
PP  
XX  
XX 28-AUG-2001; 2001WO-EP009892.  
PR  
XX  
XX (FARB ) BAYER AG.  
PA  
XX  
XX Tietjen K, Weidler M;  
PI  
XX WPI; 2002-269010/31.  
DR  
XX  
XX Identifying plant target proteins for herbicidally active compounds,  
PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
PT organisms.  
PT  
XX  
XX Claim 5; SEQ ID NO 3015; 261pp + Sequence Listing; English.  
PS  
XX  
XX The invention relates to identifying target proteins (ABB90790-ABB94016)  
CC for herbicidally active compounds, comprising aligning and comparing  
CC nucleic acid or amino acid sequences from plant with nucleic acid or  
CC amino acid sequences from non-plant organisms using suitable search  
CC parameters, where plant sequences having an E-value greater by a factor  
CC of 3 than the E-value of most similar non-plant sequences are selected.  
CC The polypeptides or nucleic acids encoding them are useful for  
CC identifying modulators. The identified modulators are useful as  
CC herbicides  
CC  
XX  
SQ Sequence 1036 AA;  
Query Match 18.7%; Score 258.5; DB 5; Length 1036;  
Best Local Similarity 30.7%; Pred. No. 1.6e-13;  
Matches 89; Conservative 50; Mismatches 98; Indels 53; Gaps 14;  
Qy 10 MLAVELGFIQVAGATSWCHDDHLALGALNLSGKGAVRLRAAW-SGASCCSWEGVG 68  
Db 5 LLLVFF-----VGSSVSQPCHPNDLSALRELALGNKSVTE---SWLNGSRCCWDGVF 56  
Qy 69 CE--TAGGRVVALRPLKRGIGGIIPSGIGELDLRLVLDLSGNSLVGEVPKSL----QIRL 122  
Db 57 CEGSDVSGRVTKLVLPKGLGEGVSKSLGELTELRLVLDLSRNQLKGEVPAEISKLEQLQV 116  
Qy 123 KSLTTDSQSLGMSGI-NMLLHVSRRTLDPEPTIGTNNVSG-----SGSNVVS 173  
Db 117 LDL---SHNLLSGVLGVSGKLQSLNTSSNSLSGLSDVGVPGLVMLNVSNLFE 173  
Qy 174 NDNTV---SGNNHVSNNVTGSDNTVGSNVHVSCTKHIVTDNNNVSG----- 224  
Db 174 EIHPELCSSGGIOVLDSNRLVGNLD---GLYNCCKSIQQLHIDSNRLTQQLPDYLY 229  
Qy 225 -----NDNNVSGSFHTVSGHNTVSGSNNTVSGSNHVSNGNK---VVTD 266  
Db 230 SIRELEQLSLGNY--LSGE--LSKNLSNLGLKSLISENRFSDVIPD 274

Search completed: May 9, 2005, 20:20:40  
Job time : 78 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 22:28:53 ; Search time 3912 Seconds  
(without alignments)  
9535.528 Million cell updates/sec

Title: US-10-657-852A-3  
Perfect score: 980  
Sequence: 1 gcttcattcccaatcaagt.....cccttacataaaaaaaa 980

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_hic:\*
- 4: gb\_est3:\*
- 5: gb\_est4:\*
- 6: gb\_est5:\*
- 7: gb\_est6:\*
- 8: gb\_gss1:\*
- 9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 372.8 | 38.0        | 767    | 4     | BJ451048    |
| C 2        | 372.4 | 38.0        | 746    | 4     | BJ458554    |
| C 3        | 370.2 | 37.8        | 686    | 4     | BJ461731    |
| C 4        | 370.2 | 37.8        | 699    | 4     | BJ461803    |
| C 5        | 366.6 | 37.4        | 687    | 4     | BJ459099    |
| C 6        | 365.6 | 37.3        | 693    | 4     | BJ457564    |
| C 7        | 365.6 | 37.3        | 696    | 4     | BJ461908    |
| C 8        | 362.4 | 37.0        | 588    | 2     | BE490074    |
| 9          | 362.2 | 37.0        | 632    | 4     | BJ453251    |
| 10         | 359.4 | 36.7        | 671    | 4     | BJ454271    |
| C 11       | 357.6 | 36.5        | 684    | 4     | BJ462238    |
| C 12       | 357.2 | 36.4        | 609    | 4     | BJ450137    |
| C 13       | 357.2 | 36.4        | 689    | 4     | BJ458121    |
| C 14       | 352   | 35.9        | 692    | 2     | BF474043    |
| C 15       | 348.2 | 35.5        | 703    | 2     | BE705098    |
| 16         | 347   | 35.4        | 625    | 4     | BJ448772    |
| 17         | 347   | 35.4        | 726    | 4     | BJ451602    |
| 18         | 347   | 35.4        | 726    | 4     | BJ454431    |
| C 19       | 346.4 | 35.3        | 830    | 2     | BE705403    |
| C 20       | 344.4 | 35.1        | 663    | 4     | BJ461352    |
| C 21       | 340.4 | 34.7        | 830    | 2     | BE705439    |
| C 22       | 335.6 | 34.2        | 1115   | 7     | CK214468    |
| 23         | 332.6 | 33.9        | 841    | 4     | BJ224369    |
| 24         | 331.6 | 33.8        | 851    | 4     | BJ448689    |

|      |       |      |     |   |          |
|------|-------|------|-----|---|----------|
| C 25 | 328.8 | 33.6 | 675 | 4 | BJ460799 |
| C 26 | 325.4 | 33.2 | 724 | 4 | BJ456413 |
| C 27 | 323   | 33.0 | 620 | 4 | BJ447068 |
| 28   | 322.2 | 32.9 | 589 | 1 | AV909087 |
| 29   | 318.6 | 32.5 | 862 | 4 | BJ454200 |
| C 30 | 316.2 | 32.3 | 818 | 7 | CK197682 |
| C 31 | 310.8 | 31.7 | 775 | 2 | BE705684 |
| C 32 | 301.6 | 30.8 | 709 | 4 | BJ453816 |
| C 33 | 296.8 | 30.3 | 648 | 4 | BJ300903 |
| C 34 | 296   | 30.2 | 574 | 1 | AV910970 |
| C 35 | 294.6 | 30.1 | 914 | 7 | CK156167 |
| C 36 | 294   | 30.0 | 612 | 1 | AV911379 |
| C 37 | 293.6 | 30.0 | 581 | 4 | BM376553 |
| 38   | 292.8 | 29.9 | 588 | 4 | BJ453528 |
| 39   | 285   | 29.1 | 480 | 1 | AJ460325 |
| 40   | 277.8 | 28.3 | 622 | 1 | AU251218 |
| C 41 | 275.8 | 28.1 | 585 | 4 | BI479842 |
| C 42 | 275   | 28.1 | 841 | 7 | CK196896 |
| C 43 | 263.2 | 26.9 | 532 | 2 | BF200590 |
| 44   | 255.6 | 26.1 | 420 | 1 | AJ460320 |
| 45   | 255.6 | 26.1 | 420 | 1 | AJ460322 |

## ALIGNMENTS

RESULT 1  
BJ451048 767 bp mRNA linear EST 23-MAY-2002  
LOCUS BJ451048 K. Sato unpublished cDNA library, cv. Akashinriki  
DEFINITION vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone  
baak32h16 5', mRNA sequence.  
ACCESSION BJ451048.1 GI:21129647  
VERSION BJ451048  
KEYWORDS EST  
SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.  
1 (bases 1 to 767)  
Sato, K., Saisho, D. and Takeda, K.  
Barley EST sequencing project in NIG and Okayama Univ  
Unpublished (2002)  
JOURNAL  
COMMENT Contact: Tadaasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers

## FEATURES

source  
1..767  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="Akashinriki"  
/sub\_species="vulgar"  
/db\_xref="taxon:112509"  
/clone="baak32h16"  
/tissue\_type="leaves"  
/dev\_stage="vegetative stage"  
/clone\_lib="K. Sato unpublished cDNA library, cv.  
Akashinriki vegetative stage leaves"

## ORIGIN

Query Match 38.0%; Score 372.8; DB 4; Length 767;  
Best Local Similarity 73.5%; Pred. No. 4.2e-100;  
Matches 524; Conservative 0; Mismatches 168; Indels 21; Gaps 3;  
Qy 233 GTGGGATCGAACAGCGCCGCGTGGCGCTTCCCGCCGCGGCGCTT 292  
Db 27 GNAGCATCTTGGCAGCGCTCGCATGTTGGAGGAGCTCAACCTTCCCAACAGACTG 86

```

QY 293 GGAGGATATCCATCTGCTGATTTGGAGCTTGATCACCCTTCGCTATTTGGATCTCTCG 352
Db 87 GTTGGACCATCCCGTCAATGATTTGGAGCTTGACACCTTTACTTGGATCTTTTCG 146
QY 353 GGTAAATCATTGTTGGGAGGTACCAAAAAGTTTGCAGATACGGCTCAAGAGCCTCAC 412
Db 147 GATAATCATTGTTGGGAGGTACCAAAAGTTT---GATACGGCTCAAGGGCTTCGCC 203
QY 413 ACTGACAGCCAGTCACTCGGTATGGTTCATTAACATGCTATTGCAATGT---GAGCAGT 469
Db 204 ATCGCTGCTGTTCAATCAGGTATGATTTTACTAATGATGCTATGATGTGGAGCCTAAC 263
QY 470 AGAAGACGCTCGATGAAGAACCAATACATATACAGGACCAACAATAGTTGGATCA 529
Db 264 AGAAGATGCTCGAGAACCAATACATATCTGGAGCAACAACACTGTGCAATCT 323
QY 530 GGGAGCAACAATGTTTTCGGGAATGACAACACGGTCTGATCTGCGGAATPAAACAACAT 589
Db 324 GGGAGCAACAATGTTTTCGGGAACGACAACACTGTCTATCCGGGAACAACAACAT 383
QY 590 GTGTCTGGAGCAACAACACTGTTGTAATCTGGAAGTGACAATCTGATGTGGTAGCAAC 649
Db 384 GTGGCTGGAGCAACAACACTATCTGTAACCGGGAACGACAATACCGTAACCTGGTAGCAAC 443
QY 650 CATGCTGATCAGGACCAACATATCTGTAATCTGTAACCAATATGTTGTAATCGGGAAC 709
Db 444 CATGCTGATCTGGGACCAACAATCTGTAATCTGTAACCAACAATGCGGTATCCGGGAAT 503
QY 710 GACAATATGCTCTGGAAGCTTCATATCTGATCAGGGAGCAACAATACCGTATCCGGG 769
Db 504 GACAATATGATCTGGGAGTTTCCATACCGTATCCGGGAACCAATCTGTAATCTGGG 563
QY 770 AGCAACAATCTGATCTGGGAGCAACAATCTGTAATCTGGGAGCAACAATGCTGTAACA 829
Db 564 ACCAACAACACTGATCTGGGAGCAACAATCTGTAATCTGGGAGCAACAATGCTGTAAGA 623
QY 830 GATGTTAATAT-----TCTGTAGTGCAGGATGCTTCCATCTTCCAA 874
Db 624 GATGATGATTTGTAAGTGGAGTGTCCATCTTCGTGAGGAGCTCACCCTTGTGTGCGA 683
QY 875 GTTCAGTGTAGTTCACATCAATAGAGGAGCAACAATCAGTTATGTAACCTCA 927
Db 684 GTTCGTTAGTCTACATCACTTGTGGGCAACATGCTGTTATGTAACCTCA 736

```

```

RESULT 2
BJ458554/c
LOCUS
DEFINITION
vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
baak32h16 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 746)
Sato, K., Saieho, D. and Takeda, K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..746
/organism="Hordeum vulgare subsp. vulgare"
FEATURES
source

```

```

/mol_type="mRNA"
/cultivar="Akashinriki"
/sub_species="vulgare"
/db_xref="taxon:1112509"
/clone="baak32h16"
/tissue_type="leaves"
/dev_stage="vegetative stage"
/clone_lib="K. Sato unpublished cDNA library, cv.
Akashinriki vegetative stage leaves"

ORIGIN

Query Match      38.0%; Score 372.4; DB 4; Length 746;
Best Local Similarity 73.7%; Pred. No. 5.5e-100;
Matches 523; Conservative 0; Mismatches 166; Indels 21; Gaps 3;

QY 236 GGATGCAAAACAGCAAGCGCGCGTCTGGTGGTGGAGAGCTCAACCTTGCACAACAAGAGCTTT 295
Db 717 GCATCCTTGGCAGCCTCGCATGGTGGAGAGCTCAACCTTGCACAACAAGAGCTTT 658
QY 296 GGGATCATCCATCGTCTGATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 355
Db 657 GGCACCATCCCGTCATGGATTTGGGAGCTTGAACCACTTTACTTACTTGGATCTTTCGAT 598
QY 356 AATTCAATGTTGGGAGGTACCAAAAAGTTTGCAGATACGGCTCAAGAGCCTCACCACT 415
Db 597 AATTCAATGTTGGGAGGTACCAAAAAGTTT---GATACGGCTCAAGGGCTTCGCCATC 541
QY 416 GACAGCAGTCACTCGGTATGGGTTCCATTAACATGCTATTTGATGT---GAGCAGTAGA 472
Db 540 GCTGCTGTTTCATCAGGTATGATTTTAACTAATGCTTGTATGGAGCTTAAACAGA 481
QY 473 AGAAGCTCGATGAAGAACCAATACAAATACAGGAGCAACAATAGTGTGGATCAGGG 532
Db 480 AGAATGCTCGAGCAACAACAATACAAATACAAATCTGGGAGCAACAATCTGATCTGGG 421
QY 533 AGCAACAATGTTGTTCCGGGAATGACAACAACGGTCTGATCTGGGAATAACAACATGTG 592
Db 420 AGCAACAATGTTGTTCTGGGAACGACAACAACATCTCATATCCGGGAACAACAACAATGTG 361
QY 593 TCTGGAGCAACAACACTGTTGTAAGTGGAGTGGATGACATCTAGTGTGGTAGCAACCAT 652
Db 360 GCTGGAGCAACAACACTGTAACCGGGAACGACAATACCGTAACTGGTAGCAACCAT 301
QY 653 GTCGTATCAGGACAAAGCATATTGTTACTGATAACAATATGTTGATCTCCGGGAACGAC 712
Db 300 GTCGTATCTGGGACAACAATCTGTAACCTGACAACAACAATGCGGTATCCGGGAATGAC 241
QY 713 AATAATGTTCTGGAAGCTTCCATATCTGATCAGGGAGGACAAATACCGTATCCGGGAGC 772
Db 240 AATAATGTTCTGGAGGTTTCCATACCGTATCCGGGAAGCCACAATCTGATCTGGGAGC 181
QY 773 AACAACTATGATCCGGAGCAACCATATCTGATCTGGGAGCAACAAGTCTGTAACAGAT 832
Db 180 AACAACTATGATCTGGGACAACAACATGTCGTATCTGGGAGCAACAAGTCTGTAACAGAT 121
QY 833 GGTAAATAT-----TCTGTAGTGTGAGGATTTGCTTCCATCTTCCCAAGTT 877
Db 120 GCATGATTTGTAAGTGGAGTGTCCATCTTCGTGACGGAGCTCACCCTTGTGTCCGAGTT 61
QY 878 CAGTGTAGTTCACATCAATAGATGGAGACAATCACCGTTATGTAACCTCA 927
Db 60 CGGTGTAGTTCACATCACTTGTGGGCGCAATCTGTTATGTAACCTCA 11

RESULT 3
BJ461731/c
LOCUS
DEFINITION
vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
baak46i20 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS

```

```

SOURCE      Hordeum vulgare subsp. vulgare
ORGANISM    Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Poaceae; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 686)
AUTHORS    Sato,K., Saisho,D. and Takeda,K.
TITLE      Barley EST sequencing project in NIG and Okayama Univ
JOURNAL    Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.

FEATURES   Location/Qualifiers
            source
            1..686
            /organism="Hordeum vulgare subsp. vulgare"
            /mol_type="mRNA"
            /cultivar="Akashinriki"
            /sub_species="vulgare"
            /db_xref="taxon:112509"
            /clone="baak46120"
            /tissue_type="leaves"
            /dev_stage="vegetative stage"
            /clone_lib="K. Sato unpublished cdna library, cv.
            Akashinriki vegetative stage leaves"

ORIGIN
Query Match      37.8%; Score 370.2; DB 4; Length 686;
Best Local Similarity 75.6%; Pred. No. 2.5e-99;
Matches 507; Conservative 0; Mismatches 143; Indels 21; Gaps 3;

QY 275 CTCCCAAGCGCGCTTGGAGGATCATCCCATCGTTCGATGGTGGAGCTTGATCACTT 334
DB 684 CTTGCCAACACAGACTGGTTGGCACCATCCCGTCATGATGGCGAGCTTGACCACTT 625
QY 335 CGCTATTTCGATCTCTCGGTAATTCATTTGGTGGGAGTACCAAAAGTTTGAGATA 394
DB 624 TACTACTTGGATCTTTCCGTAATTAATTCATTTGGTGGGAGTACCAAAAGTTT---GATA 568
QY 395 CGCTCAAGAGCGCTCAACCTGAGCAGCTGCTGATGAGTATGCTTCCATTAACATGCTA 454
DB 567 CGCTCAAGGCTTCGCCATCGCTGCTTCATCAGGTATGATTTTACTAATGCTCA 508
QY 455 TTGCATGT---GAGCAGTAGAAGAACGCTGATGAGAACCAAAATACATATCAGGAGC 511
DB 507 TTGTATGTGGAGCTTAACAGAGAAATGCTCGACGAAACCAAAATACATATCTGGAGC 448
QY 512 AACAAATAGTTTGGATCAGGAGCAACAATGTTTTCGGGAGTACCAACACGCTCGTA 571
DB 447 AACAAACATGTTCAGATCTGGAGCAACAATGTTTTCGGGAGTACCAACATCTGTCTA 388
QY 572 TCTGGGAATAACCAACCATGTGTCTGGGAGCAACAACACTGTTGTAATCTGGAAGTGACAAT 631
DB 387 TCCGGGACCAACAACATGTGGCTGGGAGCAACAACACTATCTGAACCGGAGCAGCAAT 328
QY 632 ACTGTATGTTGATGACCAACATGCTGATCAGGAGCAACAAGCATATTTGTTACTGTATCAAT 691
DB 327 ACCGTAATGTTGATGACCAACATGCTGATCTGGGAGCAACAACATATCTGTAACCAAC 268
QY 692 AATGTTGTATCCGGGAGCAACAATATGCTGCTGGAGCTTCCATCTGATCAGGAGGAG 751
DB 267 AATGCGGTATCCGGGAGTACCAATATGATCTGCGGAGTTTCATACCGGTATCCGAAGC 208
QY 752 CACAATACCGGTATCCGGGAGCAACAATATCTGATCTGGGAGCAACCATATCTGATCTGGG 811
DB 207 CACAATATCTGATCTGGGAGCAACAACATCTGATCTGGGAGCAACCATATCTGATCTGGG 148
QY 812 AGCAACAAAGTCGTAAACAGATGTTTAATAT-----TCTGTAGTGTGAGGA 856
DB 147 AGCAACAAAGTCGTAGGAGATGATGATTTGTAAGTGGAGTGTCCATCTTCCGTGACGGA 88

```

```

QY 857 TTGCTTCCATCTTCCCAAGTTCCAGTTAGCTTACAAATCAATAGATGGAGCAATACGTT 916
DB 87 GCTCACCTTGTGTTCCGAGTTCGGTGTAGCTCACAATCACTTGGTGGGCGCAATCGTGT 28
QY 917 ATGTAACCTTCA 927
DB 27 ATGTAACCTCA 17

RESULT 4
BJ461803/c
LOCUS      BJ461803
DEFINITION Hordeum vulgare subsp. vulgare linear EST 23-MAY-2002
            vegetative stage leaves Hordeum vulgare subsp. vulgare cdna clone
            baak46n01 3', mRNA sequence.
ACCESSION  BJ461803
VERSION     BJ461803.1 GI:21140313
SOURCE      EST.
KEYWORDS    Hordeum vulgare subsp. vulgare
            Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Poaceae; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 699)
AUTHORS    Sato,K., Saisho,D. and Takeda,K.
TITLE      Barley EST sequencing project in NIG and Okayama Univ
JOURNAL    Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.

FEATURES   Location/Qualifiers
            source
            1..699
            /organism="Hordeum vulgare subsp. vulgare"
            /mol_type="mRNA"
            /cultivar="Akashinriki"
            /sub_species="vulgare"
            /db_xref="taxon:112509"
            /clone="baak46n01"
            /tissue_type="leaves"
            /dev_stage="vegetative stage"
            /clone_lib="K. Sato unpublished cdna library, cv.
            Akashinriki vegetative stage leaves"

ORIGIN
Query Match      37.8%; Score 370.2; DB 4; Length 699;
Best Local Similarity 75.6%; Pred. No. 2.5e-99;
Matches 507; Conservative 0; Mismatches 143; Indels 21; Gaps 3;

QY 275 CTCCCAAGCGCGCTTGGAGGATCATCCCATCGTTCGATGGTGGAGCTTGATCACTT 334
DB 686 CTTGCCAACACAGACTGGTTGGCACCATCCCGTCATGATGGCGAGCTTGACCACTT 627
QY 335 CGCTATTTCGATCTCTCGGTAATTCATTTGGTGGGAGTACCAAAAGTTTGAGATA 394
DB 626 TACTACTTGGATCTTTCCGTAATTAATTCATTTGGTGGGAGTACCAAAAGTTT---GATA 570
QY 395 CGCTCAAGAGCGCTCAACCTGAGCAGCTGCTGATGAGTATGCTTCCATTAACATGCTA 454
DB 569 CGCTCAAGGCTTCGCCATCGCTGCTTCATCAGGTATGATTTTACTAATGCTCA 510
QY 455 TTGCATGT---GAGCAGTAGAAGAACGCTGATGAGAACCAAAATACATATCAGGAGC 511
DB 509 TTGTATGTGGAGCTTAACAGAGAAATGCTCGACGAAACCAAAATACATATCTGGAGC 450
QY 512 AACAAATAGTTTGGATCAGGAGCAACAATGTTTTCGGGAGTACCAACACGCTCGTA 571
DB 449 AACAAACATGTTCAGATCTGGGAGCAACAATGTTTTCGGGAGTACCAACACATCTGTCTA 390

```

|                       |            |  |     |
|-----------------------|------------|--|-----|
| QY                    | 572        | TCTGGGAATAACAACCATGTGTCTGGGAGCAACAACTGTTGTAACCTGGAGTGAACAAT        | 631 |
| Db                    | 389        | TCCGGGAACAACAACATGTGGCTGGGAGCAACAACACTATCGTTAAACCGGAACGAACAAT      | 330 |
| QY                    | 632        | ACTGTAGTTTGTAGCAACCATGTCTGTATCTAGGAGCAAGCATATTTACTGATAACAAT        | 691 |
| Db                    | 329        | ACCGTAACCTGTGTAGCAACCATGTCTGTATCTGGGACAAACATATCGTAACTGACAACAAC     | 270 |
| QY                    | 692        | AATGTTGTATCTGGGAACGAACAATAATGTGTCTGGAAGCTTCCATATCTGTATCAGGGGAG     | 751 |
| Db                    | 269        | AATGCCGTATCCGGGAATGACAATAATGTATCTGGGATTTCCATACCGTATCCGGAAC         | 210 |
| QY                    | 752        | CACATACCGTATCCGGGAGCAACAATCTGTATCTGGGAGCAACATATCGTATCTGGG          | 811 |
| Db                    | 209        | CACATACCTGTATCTGGGACCAACAACATCTGTATCTGGGACCAACCATGTCTGTATCTGGG     | 150 |
| QY                    | 812        | AGCAACAAAGTCGTAACAGATGGTTAATAT-----TCTGTAGGTGCAGGA                 | 856 |
| Db                    | 149        | AGCAACAAAGTCGTAGGAGTCATGTTGTAAGTGGAGTGTCCATCTTCCGTGACGGA           | 90  |
| QY                    | 857        | TGCTTTCATCTTCCAAAGTTCAGTGTAGCTTAACTCAATAGATGAGAGCAACATCAAGTT       | 916 |
| Db                    | 89         | GCTCACCTGTGTCCGAGTTCGGTGTAGCTCAACATCACTTGGTGGGCAACATCGTGT          | 30  |
| QY                    | 917        | ATGTAACCTCA 927  |     |
| Db                    | 29         | ATGTAACCTCA 19   |     |
| RESULT 5              |            |  |     |
| BJ459099/c            |            |  |     |
| LOCUS                 |            |  |     |
| DEFINITION            | BJ459099   | 687 bp mRNA linear EST 23-MAY-2002                                 |     |
|                       | BJ459099   | K. Sato unpublished cDNA library, cv. Akashinriki                  |     |
|                       |            | vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone  |     |
| ACCESSION             |            | baak34k23 3, mRNA sequence.  |     |
| VERSION               | BJ459099   |  |     |
| KEYWORDS              | BJ459099.1 | GI:21137633  |     |
| SOURCE                |            |  |     |
| ORGANISM              |            | Hordeum vulgare subsp. vulgare                                     |     |
|                       |            | Hordeum vulgare subsp. vulgare                                     |     |
|                       |            | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; |     |
|                       |            | Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;         |     |
|                       |            | Poideae; Triticeae; Hordeum.                                       |     |
| REFERENCE             |            | 1 (bases 1 to 687)   |     |
| AUTHORS               |            | Sato,K., Saisho,D. and Takeda,K.                                   |     |
| TITLE                 |            | Barley EST sequencing project in NIG and Okayama Univ              |     |
| JOURNAL               |            | Unpublished (2002)   |     |
| COMMENT               |            | Contact: Tadasu Shin-i   |     |
|                       |            | Center For Genetic Resource Information                            |     |
|                       |            | National Institute of Genetics                                     |     |
|                       |            | 1111 Yata, Mishima, Shizuoka 411-8540, Japan                       |     |
|                       |            | Tel: 81-559-81-6856  |     |
|                       |            | Fax: 81-559-81-6855  |     |
|                       |            | Email: tshini@genes.nig.ac.jp.                                     |     |
| FEATURES              |            | Location/Qualifiers  |     |
| source                |            | 1..687   |     |
|                       |            | /organism="Hordeum vulgare subsp. vulgare"                         |     |
|                       |            | /mol_type="mRNA"   |     |
|                       |            | /cultivar="Akashinriki"  |     |
|                       |            | /sub_species="vulgare"   |     |
|                       |            | /db_xref="taxon:112509"  |     |
|                       |            | /clone="baak34k23"   |     |
|                       |            | /tissue_type="leaves"  |     |
|                       |            | /dev_stage="vegetative stage"                                      |     |
|                       |            | /clone_lib="K. Sato unpublished cDNA library, cv.                  |     |
|                       |            | Akashinriki vegetative stage leaves"                               |     |
| ORIGIN                |            |  |     |
| Query Match           |            | 37.4%; Score 366.6; DB 4; Length 687;                              |     |
| Best Local Similarity |            | 74.9%; Pred.No.3e-98;  |     |
| Matches               |            | 507; Conservative 0; Mismatches 149; Indels 21; Gaps 3;            |     |
| QY                    | 275        | CTCCCCAAGCGCGCCCTTGGAGGATCATCCCATCGTGTGAGTGTGATCATCCTT             | 334 |

```

FEATURES
  source      Location/Qualifiers
    1..693
      /organism="Hordeum vulgare subsp. vulgare"
      /mol_type="mRNA"
      /cultivar="Akashinriki"
      /sub_species="vulgare"
      /db_xref="taxon:112509"
      /clone="baak28d17"
      /tissue_type="leaves"
      /dev_stage="vegetative stage"
      /clone_lib="K. Sato unpublished cdna library, cv.
      Akashinriki vegetative stage leaves"

ORIGIN
Query Match      37.3%; Score 365.6; DB 4; Length 693;
Best Local Similarity 75.3%; Pred. No. 5.9e-98;
Matches 503; Conservative 0; Mismatches 144; Indels 21; Gaps 3;

QY 275 CTCCTCAAGCGGCGCTTGGAGGATCATCCCATCGTATGGTGGAGCTTGATCACTT 334
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.685 CTTGCCAACACAGACTGGTTGGCACCATCCCGTCATGGATTGGCGAGCTTGACCACCTT 626
QY 335 CGCTATTTCGATCTCTCGGTAAATTCATTTGGTGGGAGTACCAAAAAGTTGCAGATA 394
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.625 TACTACTTGGATCTTTCCGATAATTCATTTGGTGGGAGTACCAAAAAGTTT---GATA 569
QY 395 CGGCTCAAGAGCCTCACCACTGACGACGAGTCACTCGGTATGGTTTCCATTAACATGCTA 454
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.568 CGGCTCAAGGCTTCGCCATCGCTGCTTCATCAGGTATGATTTTACTAACATGCCA 509
QY 455 TTGCATGT---GAGCAGTAGAAGACGCTCGATGAAGAACCAAAATACATATCAGGGAGC 511
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.508 TTGTATGTGGAGCCTAACAGAAATGCTCGACGAACAACCAATACATATCTCGGAGC 449
QY 512 AACATAGTTGGTATCAGGAGCAACAATTTGTTTCCGGGAATCACAACCGTTCGTA 571
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.448 AACAACTGTTCAGATCTGGGAGCAACAATTTGTTTCTGGGAACGACAACTGTCTA 389
QY 572 TCTGGGAATAACAAACCATGTGTCTGGGAGCAACAACACTGTTGTAACCTGGAAGTGACAAT 631
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.388 TCCGGGACAAACAATGTGGCTGGGAGCAACAACACTATCGTAACCGGGACGACAT 329
QY 632 ACTGTAGTTGGTAGCAACCATGCTGATCAGGGAACAAGCATATGTTTACTGATAACAAT 691
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.328 ACCGTAACCTGTTAGCAACCATGCTGATCTGGGGAACAACATATCTGTAACCTGACAACAAC 269
QY 692 AATGTTGTATCCGGGAGCAACAATATGCTCTGGAGCTTCCATCTGATATCAGGGAG 751
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.268 AATGCCGTATCCGGGAATGCAATATGATCTGGGAGTTTCCATACCGTATCCGGAGC 209
QY 752 CACAATACCGTATCCGGGAGCAACAATATGATCTGGGAGCAACAACATATCGTATCTGGG 811
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.208 CACATATCTGTATCTGGGACCAACAACATGATCTGGGAACAACCATGTCGTAATCTGGG 149
QY 812 AGCAACAAGTCTGTAACAGATGTTAATAT-----TCTGTAGTGCAGGA 856
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.148 AGCAACAAGTCTGTAAGATGATGATTTGTAAGTGGAGTGCCATCTTCGGTGACGGA 89
QY 857 TTGCTTCCATCTTCCCAAGTTAGTGTAGCTTCAATATCAATAGATGAGACAAATACAGTT 916
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.88 GCTCAGCTGTTTGTCCGAGTTCGGTGTAGTCAATATCACTTGGTGGGGCCAAATCGTGT 29
QY 917 ATGTAACCT 924
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.28 ATGAACT 21

RESULT 7
BJ461908/c 696 bp mRNA linear EST 23-MAY-2002
LOCUS BJ461908 K. Sato unpublished cdna library, cv. Akashinriki
DEFINITION vegetative stage leaves Hordeum vulgare subsp. vulgare cdna clone
baak21d01 3', mRNA sequence.

```

```

ACCESSION BJ461908
VERSION BJ461908.1 GI:21140417
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
1 (bases 1 to 696)
Sato, K., Saisho, D. and Takeda, K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadabu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
  1..696
    /organism="Hordeum vulgare subsp. vulgare"
    /mol_type="mRNA"
    /cultivar="Akashinriki"
    /sub_species="vulgare"
    /db_xref="taxon:112509"
    /clone="baak21d01"
    /tissue_type="leaves"
    /dev_stage="vegetative stage"
    /clone_lib="K. Sato unpublished cdna library, cv.
    Akashinriki vegetative stage leaves"

ORIGIN
Query Match      37.3%; Score 365.6; DB 4; Length 696;
Best Local Similarity 75.3%; Pred. No. 5.9e-98;
Matches 503; Conservative 0; Mismatches 144; Indels 21; Gaps 3;

QY 275 CTCCTCAAGCGGCGCTTGGAGGATCATCCCATCGTATGGTGGAGCTTGATCACTT 334
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.683 CTTGCCAACACAGACTGGTTGGCACCATCCCGTCATGGATTGGCGAGCTTGACCACCTT 624
QY 335 CGCTATTTCGATCTCTCGGTAAATTCATTTGGTGGGAGTACCAAAAAGTTGCAGATA 394
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.623 TACTACTTGGATCTTTCCGATAATTCATTTGGTGGGAGTACCAAAAAGTTT---GATA 567
QY 395 CGGCTCAAGAGCCTCACCACTGACGACGAGTCACTCGGTATGGTTTCCATTAACATGCTA 454
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.566 CGGCTCAAGGCTTCGCCATCGCTGCTGTTTCTAGGATGATTTTACTAACATGCCA 507
QY 455 TTGCATGT---GAGCAGTAGAAGACGCTCGATGAAGAACCAAAATACATATCAGGGAGC 511
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.506 TTGTATGTGGAGCCTAACAGAAAGATGCTCGACGAACAACCAATACATATCTGGGAGC 447
QY 512 AACATAGTTGGTAGTACGGGAGCAACAATTTGTTTCCGGGAATGACAACACGGTTCGTA 571
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.446 AACAACTGTTCAGATCTGGGAGCAACAATTTGTTTCTGGGAACGACAACACTGTCTA 387
QY 572 TCTGGGAATAACAAACCATGTGTCTGGGAGCAACAACACTGTTGTAACCTGGAAGTGACAAT 631
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.386 TCCGGGAACAACAACATGTTGGCTGGGAGCAACAACACTATCGTAACCGGGAGACGACAAT 327
QY 632 ACTGTAGTTGGTAGCAACCATGTCTGATCAGGGAACAAGCATATTTGTTACTGATAACAAT 691
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.326 ACCGTAACCTGTTAGCAACCATGTCGTTCTGGGACAAACATATCTGTAACCTGACAACAAC 267
QY 692 AATGTTGTATCCGGGAGCAACAATATGATGTTCTGGAGCTTCCATCTGATATCAGGGAG 751
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.266 AATGCCGTATCCGGGAATGACAATATGATCTGGGAGTTTCCATACCGCTTCCGGAAGC 207
QY 752 CACAATACCGTATCCGGGAGCAACAATATGATCTGGGAGCAACAACACTATCGTATCTGGG 811
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
.206 CACAATACTGTATCTGGGAGCAACAACACTGTATCTGGGAACAACCAATCTGTTCTGGG 147

```

QY 812 AGCAACAAAGTCGTACACAGATGGTTAATAT-----TCTGTAGTGCAGGA 856  
 |||||  
 Db 146 AGCAACAAAGTCGTAGAGATGCATGATTTGTAAGTGGAGTGTCCATCTTCGTGACCGGA 87  
 |||||  
 QY 857 TTGCTTCCATCTTCCCAAGTTCAGTGTAGCTTACCAATCAATAGATGGAGACAAATCAAGTT 916  
 |||||  
 Db 86 GCTCACCTTGTTGTCGAGTTCGGTGTAGCTCACATCACTACCTTGGTGGGGCCAACTGTT 27  
 |||||  
 QY 917 ATGTAACCT 924  
 |||||  
 Db 26 ATGAACCT 19

BE490074 588 bp mRNA linear EST 31-JUL-2000  
 WHE0365 G07 N13ZS Wheat cold-stressed seedling cDNA library  
 LOCUS Triticum aestivum cDNA clone WHE0365\_G07\_N13, mRNA sequence.  
 DEFINITION BE490074  
 ACCESSION BE490074.1 GI:9609607  
 VERSION EST.  
 KEYWORDS Triticum aestivum (bread wheat)  
 SOURCE Triticum aestivum  
 ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.  
 1 (bases 1 to 588)

AUTHORS Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D.,  
 Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,  
 Seaton, C.L. and Tong, J.C.  
 TITLE The structure and function of the expressed portion of the wheat  
 JOURNAL genomes - Cold-stressed seedling cDNA library  
 COMMENT Unpublished (2000)

CONTACT: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818

Email: oanderson@pw.usda.gov  
 Sequence have been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20  
 Seq primer: Stratagene SK primer.

## FEATURES

Location/Qualifiers  
 1..588  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Chinese Spring"  
 /db\_xref="taxon:4565"  
 /clone="WHE0365\_G07\_N13"  
 /tissue\_type="Seedling"  
 /dev\_stage="Five-day old seedling"  
 /lab\_host="E. coli SOLR"  
 /clone\_lib="Wheat cold-stressed seedling cDNA library"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
 Site 1: EcoRI; Site 2: XhoI; Seeds were  
 surface-sterilized, germinated and grown aseptically in  
 the dark at room temperature on filter paper with water,  
 nystatin and cefotaxime in covered crystallization  
 dishes. Five-day old seedlings were transferred to 5 C  
 cold room and kept for 48 hr. The tissue, total RNA, and  
 poly(A) RNA were prepared, a cDNA library was made, and  
 the cDNA clones were in vivo excised to give pBluescript  
 phagemids in the TJ Close lab (Choi, Close, Fenton) at  
 the University of California, Riverside. Plasmid DNA  
 preparations and DNA sequencing were performed in the OD  
 Anderson lab (all other authors)."

## ORIGIN

Query Match 37.0%; Score 362.4; DB 2; Length 588;  
 Best Local Similarity 79.5%; Pred. No. 5.2e-97;  
 Matches 455; Conservative 0; Mismatches 111; Indels 6; Gaps 2;

QY 257 CGCGTGTGGCGGTTCGGGCTCCCCAAGCGCGGCCCTTGGAGGGATCATCCCATCCTCGATT 316  
 |||||  
 Db 20 CGGCTAGAGGAGCTCAACCTTGGCCAAACAACAACTGGTCGGTACCATCCCATCGTGATT 79  
 |||||  
 QY 317 GGTGAGCTTGATCACCTTCGCTATTTCGATCTCTCGGTTAATTCATTGGTTGGGAGGTA 376  
 |||||  
 Db 80 GCGAGCTTGATCACCTTTGCTACTTGGATCTGCGGATTAATTCATTGGTTGGGAGGTA 139  
 |||||  
 QY 377 CCAAAAAGTTTGCAGATACGGCTCAAGAGCCTCACCACTGACAGCCAGTCACTCGGTATG 436  
 |||||  
 Db 140 CCCAAGAGTTT---GATAAGGCTCAAGGTTCTCGTCATCGTTGGTCTCACTAGGTATG 196  
 |||||  
 QY 437 GGTCCATTAACTGCTATTTCATGTGA---GCAGTAGAAGAACGCTCGATGAAGAACA 493  
 |||||  
 Db 197 GTTTTACGAACATGCTCCATTTGATGTGAAGCGTAATAGAAGAACACTCGAGCAACAACA 256  
 |||||  
 QY 494 AATCAATATATCAGGAGCCCAACAATAGTGTGGATCAGGAGCAACAATTTGTTTCGGG 553  
 |||||  
 Db 257 ANTACATATCTGGAGCAACACACTGTGATCTCGAAGCACCAAGTTGTTTCTGGG 316  
 |||||  
 QY 554 AATCAACACAGGTCGTATCTGGGAATAACAACCATGTGTCTGGGAGCAACAACACTGTT 613  
 |||||  
 Db 317 AATGACAACACTGTATATCGGGGAATAACAACAATGTGGCTGTAGCAACAACACTGTC 376  
 |||||  
 QY 614 GTAACCTGAGTGAACATCTAGTTGGTAGCAACCATCTCGTATCAGGAGCAAAAGCAT 673  
 |||||  
 Db 377 GTAAACCGGGAAACAACAATACCGTAACCTGGTAGCAACCATATTTGTTATCTGGGAGCAACA 436  
 |||||  
 QY 674 ATTGTTACTGATACAATAATTTGTATCCGGGAACCAACAATAATGTCTGGAAGCTTC 733  
 |||||  
 Db 437 ATCGTGACTGACATAACATGCGGTATCCGGNATGACAAATATGTATCTGGAGGCTTC 496  
 |||||  
 QY 734 CATACTGTATCAGGGAGGCAATACCGTATCCGGGAGCAACAATCTGTATCCGGAGC 793  
 |||||  
 Db 497 CATACCTATCCGGAGGCCAATACTGTATCTGGGACCAACAACACTGTATCCGGAAGC 556  
 |||||  
 QY 794 AACCATATCTATCTGGGAGCAACAAGTCGT 825  
 |||||  
 Db 557 AACCATGTGTATCTGGGAGCAACAAGTCGT 588  
 |||||

## RESULT 9

BJ453251

## LOCUS

DEFINITION

BJ453251 632 bp mRNA linear EST 23-MAY-2002  
 vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone  
 baak42e04 5', mRNA sequence.

## ACCESSION

BJ453251

## VERSION

BJ453251.1

## KEYWORDS

GI:21131831

## SOURCE

Est.

## ORGANISM

Hordeum vulgare subsp. vulgare

Hordeum vulgare subsp. vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Hordeum.

1 (bases 1 to 632)

Sato, K., Saisho, D. and Takeda, K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..632

/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/cultivar="Akashinriki"

/sub\_species="vulgare"

/db\_xref="taxon:112509"

```

/clone="baak42e04"
/tissue_type="leaves"
/dev_stage="vegetative stage"
/clone_lib="K. Sato unpublished cDNA library, cv.
Akashinriki vegetative stage leaves"

ORIGIN

Query Match      37.0%; Score 362.2; DB 4; Length 632;
Best Local Similarity 77.1%; Pred. No. 6e-97;
Matches 468; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 241 CGAAACAGCAAGCGCGCGCTGCGGCTGCGGCTCCCAAGCGCGGCTTGGAGGAT 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1  CGAGGAGCGGCTCGCATGGTGGAGGAGCTCAACCTTGCCAAACACAGACTGGTTGGCAC 60

QY 301 CATCCCATCGTCATTTGGTGAGCTTCATCACCTTCGCTATTTGGATCTCTCGGGTAATTC 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 CATCCCGTCATGATTTGGCGAGCTTGACACCTTTACTACTTGGATCTTGGATTAATTC 120

QY 361 ATTTGGTTGGGAGGATCAAAAAGTTTGCAGATACCGGCTCAAGAGCCTCAACCTGACAG 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 ATTTGGTTGGCGAGGTACCAAGAGTTT---GATACGGCTCAAGGGCTTCGCCATCGCTGG 177

QY 421 CGAGTCACCTCGGTATGGTTCCATTAACATGCTATTTCATGT---GAGCAGTAGAAGAC 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 178 TCGTTTCATCAGGTATGATTTTACTAATCATGCTGCTATTTGATGAGGCTTAACAGAGAT 237

QY 478 GCTCGATGAAGAACCAATACATATATCAGGGACCAACATAGTGTGGATCAGGGAGCAA 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 238 GCTCGAGAACCAACCAATACATATATCTGGGAGCAACACACTGTGAGATCTGGAGCAC 297

QY 538 CAATGTTGTTCCGGGAATGACAAACGCTGCTATCTGGGAATPAACAACCATGTGCTGG 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 298 CAATGTTGTTCCGGGAACGACAAACACTGTCTATCCGGGAACAAACAATGTGGCTGG 357

QY 598 GAGCAACACACTGTTGTAAGTGAAGTGAACATCTGATGTTGGTAGCAACCATGTGCT 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 358 GAGCAACACACTGTGTAACCGGGAACGCAATACCGTAACCTGTTAGCAACCATGTGCT 417

QY 658 ATCAGGGACAAAGCATATTTGTAACATAATATTTGTTATCCGGGAACCAACATAA 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 418 ATCTGGGACAAACATATGTAATCTGACAAACATGCGGTATCCGGGAATGACATTA 477

QY 718 TGTGCTCGGAAGTTCCTACTGTATCAGGGAGCACAATACCGTATCCGGGAGCAACAA 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 478 TGTATCTGGGAGTTTCATACCGTATCCGGAAGCCACATCTGTATCTGGGACCAACAA 537

QY 778 TACTGTATCCGGGAGCAACCATATGTAATCTGGGAGCAACAAAGTGTGAACAGATGTTA 837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 538 CACTGTATCTGGGAACAAACCATGCTGTAATCTGGGAGCAACAAAGTGTAGGAGATCATG 597

QY 838 ATATTCT 844
DB 598 ATTTGT 604

RESULT 10
BJ454271
LOCUS
DEFINITION
    BJ454271 K. Sato unpublished cDNA library, cv. Akashinriki
    vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
    baak46n01 5', mRNA sequence.
ACCESSION
    BJ454271
VERSION
    BJ454271.1 GI:21132845
KEYWORDS
    Hordeum vulgare subsp. vulgare
SOURCE
    Hordeum vulgare subsp. vulgare
ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
    Pooideae; Triticaceae; Hordeum.
REFERENCE
    1 (bases 1 to 671)
    Sato,K., Saisho,D. and Takeda,K.
    Barley EST sequencing project in NIG and Okayama Univ

```

```

Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
    source
    1..671
    /organism="Hordeum vulgare subsp. vulgare"
    /mol_type="mRNA"
    /cultivar="Akashinriki"
    /sub_species="vulgare"
    /db_xref="taxon:112509"
    /clone="baak46n01"
    /tissue_type="leaves"
    /dev_stage="vegetative stage"
    /clone_lib="K. Sato unpublished cDNA library, cv.
    Akashinriki vegetative stage leaves"

```

```

ORIGIN

Query Match      36.7%; Score 359.4; DB 4; Length 671;
Best Local Similarity 76.7%; Pred. No. 4.2e-96;
Matches 467; Conservative 0; Mismatches 136; Indels 6; Gaps 2;

QY 236 GGATCGAAACAGCAAGCGCGCGCTGCGGCTTCCCGCTCCCAAGCGCGGCTTGGGA 295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 53 GCATCTTTGGCAGGCTCGCATGGTTGGAGGAGCTCAACCTTGCCAAACAGACTGGTT 112

QY 296 GGGATCATCCCATCGTTCGATTTGGTAGCTTGATCACCTTCGCTTATTTGGATCTCTCGGGT 355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 113 GGCACCATCCCGTCATGGATTGGCGAGCTTGACCACCTTTACTACTTGGATCTTTCGGAT 172

QY 356 AATTCAATGTTGGGGAGGTACCAAAAGTTTGCAGATACGCTCAAGAGCCTCACCACCT 415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 173 AATTCAATGTTGGGGAGGTACCCAAAGTTT---GATACGGCTCAAGGGCTTCGCCATC 229

QY 416 GACAGCAGCTACTCGGTATGGGTTTCCATTAACATGCTATTGTCATGT---GAGCAGTAGA 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 230 GCTGGTTCGTTTCATCAGGTATGATTTTACTAACATGCCATTTGATGTGGAGCCTAACAGA 289

QY 473 AGAAGCTCGATGAGAACAACAAATACATATCAGGGACCAACATAGTGTGGATCAGGG 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 290 AGAATGCTCGAGGAACAACAAATACATATCTGGGAGCAACAACTGTGAGATCTGGG 349

QY 533 AGCAACAATGTTGTTTCCGGGAATGACACACGCTGATCTGGGAATACAAACCATGTG 592
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 350 AGCACCATGTTGTTTCTGGGAACGACAACTGTCTATCTGGGGAACAAACAATGTG 409

QY 593 TCTGGGAGCAACAACACTCTGTTAACTGGAAGTGACAATACTGTAGTGTGTAGCAACCAT 652
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 410 GCTGGGAGCAACAACACTATCGTAACCGGGAACGACAATACCGTAACTGTTAGCAACCAT 469

QY 653 GTCGTATCAGGGAACAAGCATATTTGTAATGATAACAATAATGTTTATCTCGGGGAACGAC 712
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 470 GTCGTATCTGGGGAACAACATATCGTAACCTGACAACAACAACTGCTATCTCGGGGAATGAC 529

QY 713 AATAATGCTCTGGAGCTTCCATCTGATCAGGGGAGCAACAATACCGTATCTCGGGGAGC 772
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 530 AATAATGTAATCTGGGAGTTTCCATACCGGTATCCGGAAGCAACAATACTGTATCTCGGGACC 589

QY 773 AACAACTACTGTATCCGGGAGCAACCATATCTGATCTGGGAGCAACAAGTGTGTAACAGAT 832
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 590 AACAACTACTGTATCTGGGGAACAACCATGTCGTATCTGGGAGCAACAAGTGTGAGGAT 649

QY 833 GGTAAATAT 841
DB 650 GCATGATTT 658

```

```

RESULT 11
BJ462238/c

```

```

LOCUS      BJ462238                684 bp    mRNA    linear    EST 23-MAY-2002
DEFINITION      BJ462238 K. Sato unpublished cDNA library, cv. Akashinriki
                vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
                baak24g16 3', mRNA sequence.
ACCESSION      BJ462238
VERSION        BJ462238.1    GI:21140745
KEYWORDS       EST.
SOURCE         Hordeum vulgare subsp. vulgare
ORGANISM       Hordeum vulgare subsp. vulgare
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
               Poideae; Triticeae; Hordeum.
REFERENCE      1 (bases 1 to 684)
AUTHORS        Sato, K., Saisho, D. and Takeda, K.
TITLE          Barley EST sequencing project in NIG and Okayama Univ
JOURNAL        Unpublished (2002)
COMMENT        Contact: Tadasu Shin-i
                Center For Genetic Resource Information
                National Institute of Genetics
                1111 Yata, Mishima, Shizuoka 411-8540, Japan
                Tel: 81-559-81-6856
                Fax: 81-559-81-6855
                Email: tshini@genes.nig.ac.jp.
FEATURES       Location/Qualifiers
               1..684
               /organism="Hordeum vulgare subsp. vulgare"
               /mol_type="mRNA"
               /cultivar="Akashinriki"
               /sub_species="vulgare"
               /db_xref="taxon:112509"
               /clone="baak24g16"
               /tissue_type="leaves"
               /dev_stage="vegetative stage"
               /clone_lib="K. Sato unpublished cDNA library, cv.
               Akashinriki vegetative stage leaves"
ORIGIN
Query Match      36.5%; Score 357.6; DB 4; Length 684;
Best Local Similarity 78.4%; Pred. No. 1.5e-95;
Matches 455; Conservative 0; Mismatches 113; Indels 6; Gaps 2;

QY 265 GCGCTGGCGGCTCCCAAGCGCGGCTTGGAGGGATCATCCATCGTCGATTTGGTGAGCT 324
DB 679 GGAGCTCAACCTTGGCAACACAGACTGGTTGGCACCATCCGTCATGGATTGGCGAGCT 620
QY 325 TGATCACCTTCGCTATTGGATCTCTCGGTAATTCATGTTGGGAGGTACCAAAAG 384
DB 619 TGACCACCTTACTACTTGGATCTTTCCGATAATTCATTGGTTGGCGAGGTACCAAGAG 560
QY 385 TTTCAGATAGCGCTCAAGAGCCTCACCACCTGCACAGCCAGTCACCTCGGTATGGTTCCAT 444
DB 559 TTT---GATACGGCTCAAGGGCTTCGCCATCGTGTGTCATCAGGTATGATTTTAC 503
QY 445 TAACATGCTATTGCATGT---GAGCAGTAGAAGAACGCTCGATGAAGAACCAAAATACAAT 501
DB 502 TAACATGCCATTGTATGTGGAGCCTTAACAGAAGAATGCTCGACGAACAACCAATACAAT 443
QY 502 ATCAGGGACCAACAATAGTTGGATCAGGAGCAACAATGTTTTCGCGGAATGACAA 561
DB 442 ATCTGGGAGCAACAACACTGTGAGATCTGGAGCACCACCAATGTTGTTTCGGGAACGACAA 383
QY 562 CACGCTCGTATCTGGGAATAACAACATGTGCTTGGGAGCAACAACACTGTTGTAACATGG 621
DB 382 CACTGTCTATACCGGGAACAACAATGTGGCTGGGAGCAACAACACTATCGTAACCGG 323
QY 622 AAGTGACAATCTGTAGTTGGTAGCAACCATGTCGTATCAGGGACAAAGCATATTTGTAC 681
DB 322 GAACGACAATACCGTAACTGTGTAGCAACCATGTGCTATCTGGGACAAACATATCGTAAC 263
QY 682 TGATAACAATAATGTTGTATCCGGGAACGACAATAATGCTGTGGAAGCTTCCATCTGT 741
DB 262 TGACAAACAACANTGCCGTATCCGGGAATGACAATAATGATATCTGGGAGTTTCCATACCGT 203

```

```

QY 742 ATCAGGGAGCACAATACCGTATCCGGAGCAACAATATCTGTATCCGGGAGCAACCATAT 801
DB 202 ATCCGGAAGCCACAATATCTGTATCTGGGACCAACAACACTGTATCTGGGAAACAACCATGT 143
QY 802 CGTATCTGGGAGCAACAAGTCGTAACAGATGGTTTAATAT 841
DB 142 CGTATCTGGGAGCAACAAGTCGTAAGAGATGCATGATTT 103

RESULT 12
BJ450137
LOCUS      BJ450137                609 bp    mRNA    linear    EST 23-MAY-2002
DEFINITION      BJ450137 K. Sato unpublished cDNA library, cv. Akashinriki
                vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
                baak28d17 5', mRNA sequence.
ACCESSION      BJ450137
VERSION        BJ450137.1    GI:21128743
KEYWORDS       EST.
SOURCE         Hordeum vulgare subsp. vulgare
ORGANISM       Hordeum vulgare subsp. vulgare
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
               Poideae; Triticeae; Hordeum.
REFERENCE      1 (bases 1 to 609)
AUTHORS        Sato, K., Saisho, D. and Takeda, K.
TITLE          Barley EST sequencing project in NIG and Okayama Univ
JOURNAL        Unpublished (2002)
COMMENT        Contact: Tadasu Shin-i
                Center For Genetic Resource Information
                National Institute of Genetics
                1111 Yata, Mishima, Shizuoka 411-8540, Japan
                Tel: 81-559-81-6856
                Fax: 81-559-81-6855
                Email: tshini@genes.nig.ac.jp.
FEATURES       Location/Qualifiers
               1..609
               /organism="Hordeum vulgare subsp. vulgare"
               /mol_type="mRNA"
               /cultivar="Akashinriki"
               /sub_species="vulgare"
               /db_xref="taxon:112509"
               /clone="baak28d17"
               /tissue_type="leaves"
               /dev_stage="vegetative stage"
               /clone_lib="K. Sato unpublished cDNA library, cv.
               Akashinriki vegetative stage leaves"
ORIGIN
Query Match      36.4%; Score 357.2; DB 4; Length 609;
Best Local Similarity 79.1%; Pred. No. 1.9e-95;
Matches 451; Conservative 0; Mismatches 113; Indels 6; Gaps 2;

QY 275 CTCCCAGGCGCGCCTTGGAGGATCATCCCATCGTCGATTTGGTGAGCTTGATCACCTT 334
DB 29 CTTCGCAACAACAGACTGGTTGGCACCATCCGTCATGGATTGGCGAGCTTGACCACTT 88
QY 335 CGCTATTGTGATCTCTCGGTAATTCATTTGGTTGGGAGGTACCAAAAAGTTTGACATA 394
DB 89 TACTACTTGGATCTTTTCGGATAATTCATTGTTGGGAGGTACCCAGAGTTT---GATA 145
QY 395 CGGCTCAAGAGCCTCACCACTGACAGCAGTCACCTCGGTATGGTTCCATTAAACATGCTA 454
DB 146 CGGCTCAAGGCTTCGCCATCGCTGTCGTTTCATCAGGTATGATTTTACTTAACATGCCA 205
QY 455 TTGCATGT---GAGCAGTAGAAGAACGCTCGATGAAGAACCAAAATACAATATCAGGAGCC 511
DB 206 TTGTATGTGGAGCCTTAACAGAAGATGCTCGACGAACAACCAATAACAATATCTGGGAGC 265
QY 512 AACATAGTTTGGATCAGGAGCAACAATGTTTTCGCGGAATGACACACAGCTGCTA 571
DB 266 AACAACTCTGAGATCTGGGAGCAACAATGTTTGTCTGGGAAACGACAACTGTGCTATA 325
QY 572 TCTGGGAATAACAACCATGTGTCTGGGAGCAACAACACTGTTGTAACGTGAAGTGACAAAT 631

```

|            |  |  |     |
|------------|--|--|-----|
| Db         | 564  | CGGCTCAAGGGCTTCGCCATCGCTGGTTCATCAGGTATGATTTTATCTAAACATGCCA     | 505 |
| Qy         | 455  | TTGCATGT---GAGCAGTAGAAGAAACGCTCGATGAAGAACCAATACAAATATCAGGGACC  | 511 |
| Db         | 504  | TTGTATGTGGAGCCTAACAGAGNATGCTCGACGAACCAACCAATATCTGGGAGC         | 445 |
| Qy         | 512  | AACAATAGTGTGGATCAGGGAGCAACAATGTTGTTTTCGGGAATGACAACAGGTGTA      | 571 |
| Db         | 444  | AACAACACTGTCTGATCTGGGAGCACCAATGTTGTTTCTGGGAACGACAACCTGTCTATA   | 385 |
| Qy         | 572  | TCTGGGAATACACCAATGTTCTGGGAGCAACAACACTGTTCTTAAGTGAAGTGACAAT     | 631 |
| Db         | 384  | TCCGGGAACCAACCAATGTTGGCTGGGAGCAACAACACTATCGTTAAACCGGGAACGACAAT | 325 |
| Qy         | 632  | ACTGTAGTGTGGTAGCAACCAATGCTATCAGGGACAAAGCATATTTGTTACTGATAACAAT  | 691 |
| Db         | 324  | ACCGTAACCTGTAGCAACCAATGCTATCTGGGGACAAACATATCGTTAACTGACAACAAC   | 265 |
| Qy         | 692  | AATGTTGTATCCGGGAACGACAATATGTTCTTGGGAAGCTTCCATATCTGATCAGGGGAG   | 751 |
| Db         | 264  | AATGCCGTATCCGGGAATGACAATAATGTTATCTGGGAGTTTCCATACCGTATCCGGAAGC  | 205 |
| Qy         | 752  | CACAATACCTGATCCGGGACCAACATCTGTTATCCGGGACCAACCATATCTGATCTGGG    | 811 |
| Db         | 204  | CACAATACTGTATCTGGGACCAACAACACTGTATCTGGGAACAACCATGCTATCTGGG     | 145 |
| Qy         | 812  | AGCAACAAAGTCTGAACAGATGGTTAATAT                                 | 841 |
| Db         | 144  | AGCAACAAAGTCTGAGGAGATGATGATTT                                  | 115 |
| RESULT 14  |  |  |     |
| BF474043   |  |  |     |
| LOCUS      | BF474043   | 692 bp mRNA linear EST 04-DEC-2000                             |     |
| DEFINITION | WHE0840_D01_H02ZS wheat vernalized crown cDNA library Triticum aestivum cDNA clone WHE0840_D01_H02, mRNA sequence.   |  |     |
| ACCESSION  | BF474043   |  |     |
| VERSION    | BF474043.1   | GI:11543225  |     |
| KEYWORDS   | EST.   |  |     |
| SOURCE     | Triticum aestivum (bread wheat)  |  |     |
| ORGANISM   | Triticum aestivum  |  |     |
| REFERENCE  | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.   |  |     |
| AUTHORS    | 1 (bases 1 to 692)<br>Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han,P.S., Haia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.   |  |     |
| TITLE      | The structure and function of the expressed portion of the wheat genomes - Vernalized crown cDNA library   |  |     |
| JOURNAL    | Unpublished (2000)   |  |     |
| COMMENT    | Contact: Olin Anderson<br>US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center<br>800 Buchanan Street, Albany, CA 94710, USA<br>Tel: 5105595773<br>Fax: 5105595818<br>Email: oanderson@pw.usda.gov<br>Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20<br>Seq primer: Stratagene SK primer.<br>Location/Qualifiers |  |     |
| FEATURES   | 1..692   |  |     |
| source     | /organism="Triticum aestivum"  |  |     |
|            | /mol_type="mRNA"   |  |     |
|            | /cultivar="Chinese Spring"   |  |     |
|            | /db_xref="taxon:4565"  |  |     |
|            | /clone="WHE0840_D01_H02"   |  |     |
|            | /tissue_type="Crown tissue of seedling"  |  |     |
|            | /dev_stage="Five-week old seedling"  |  |     |
|            | /lab_host="E. coli SOLR"   |  |     |
|            | /clone_lib="wheat vernalized crown cDNA library"   |  |     |
|            | /notes="Vector: Lambda Uni-ZAP XR, excised phagemid;   |  |     |



Db 74 GAATGATT 66

Search completed: May 11, 2005, 01:33:58  
Job time : 3917 secs

*This Page Blank (uspio)*

| Result No. | Score | Query |     | Length  | DB | ID                  | Description        |
|------------|-------|-------|-----|---------|----|---------------------|--------------------|
|            |       | Match | %   |         |    |                     |                    |
| C          | 1     | 44.8  | 4.6 | 3453    | 4  | US-10-101-464A-861  | Sequence 861, Appl |
|            | 2     | 44.8  | 4.6 | 7218    | 1  | US-08-232-463-14    | Sequence 14, Appl  |
|            | 3     | 41    | 4.2 | 408     | 4  | US-10-101-464A-212  | Sequence 212, Appl |
| C          | 4     | 40.4  | 4.1 | 396     | 4  | US-10-101-464A-221  | Sequence 221, Appl |
|            | 5     | 40.4  | 4.1 | 3381    | 3  | US-09-336-447A-6    | Sequence 6, Appl   |
|            | 6     | 40.4  | 4.1 | 3381    | 4  | US-09-952-267B-6    | Sequence 6, Appl   |
|            | 7     | 38.8  | 4.0 | 1028    | 3  | US-09-249-180-1     | Sequence 1, Appl   |
|            | 8     | 38.8  | 4.0 | 1476    | 3  | US-09-434-288-12    | Sequence 12, Appl  |
|            | 9     | 38.8  | 4.0 | 2900    | 4  | US-10-101-464A-840  | Sequence 840, Appl |
|            | 10    | 38.4  | 3.9 | 1107    | 4  | US-10-101-464A-268  | Sequence 268, Appl |
| C          | 11    | 38.4  | 3.9 | 2735    | 4  | US-10-101-464A-865  | Sequence 865, Appl |
|            | 12    | 38.2  | 3.9 | 454     | 3  | US-09-228-986-39    | Sequence 39, Appl  |
|            | 13    | 38.2  | 3.9 | 454     | 4  | US-10-101-464A-39   | Sequence 39, Appl  |
|            | 14    | 38.2  | 3.9 | 498     | 4  | US-10-101-464A-421  | Sequence 421, Appl |
|            | 15    | 38.2  | 3.9 | 505     | 4  | US-09-621-976-15639 | Sequence 15639, A  |
|            | 16    | 38.2  | 3.9 | 1590    | 4  | US-09-248-796A-2600 | Sequence 2600, App |
|            | 17    | 38.2  | 3.9 | 2691    | 4  | US-10-101-464A-837  | Sequence 837, Appl |
| C          | 18    | 38.2  | 3.9 | 23210   | 4  | US-09-596-002-17    | Sequence 17, Appl  |
|            | 19    | 38    | 3.9 | 535     | 3  | US-09-056-556-171   | Sequence 171, Appl |
|            | 20    | 38    | 3.9 | 535     | 3  | US-09-072-596-166   | Sequence 166, Appl |
|            | 21    | 38    | 3.9 | 535     | 4  | US-09-072-967-171   | Sequence 171, Appl |
|            | 22    | 38    | 3.9 | 4403765 | 3  | US-09-103-840A-2    | Sequence 2, Appl   |
|            | 23    | 38    | 3.9 | 4411529 | 3  | US-09-103-840A-1    | Sequence 1, Appl   |
|            | 24    | 37.4  | 3.8 | 4689    | 3  | US-09-105-537-34    | Sequence 34, Appl  |
| C          | 25    | 37.4  | 3.8 | 36778   | 3  | US-09-105-537-5     | Sequence 5, Appl   |
|            | 26    | 37.4  | 3.8 | 38506   | 4  | US-09-320-878-19    | Sequence 19, Appl  |
| C          | 27    | 37.4  | 3.8 | 38506   | 4  | US-09-141-908-1     | Sequence 1, Appl   |



```
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-221

Query Match
Best Local Similarity 4.1%; Score 40.4; DB 4; Length 396;
Matches 71; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 259 CQTGCTGCGCTTCCGCTCCCAAGCGCGGCTTGGAGGATCATCCCATCGTGGATTGG 318
Db 78 CCTCGAGGTCTCTGCTCAGCAGTGGCGCTTACCGGAGATCCCGGAGTCTGCTCGG 137

QY 319 TGAGCTTGATCACCTTCGCTATTGATCTCTCGGTAATTCATTGTTGGGAGGTACC 378
Db 138 GAGGCTCAAGAGGCTCACCGATTGGACCTCGCCTTCAACTACTTGGAGGGCCCATACC 197

QY 379 AA 380
Db 198 GA 199

Query Match
Best Local Similarity 4.1%; Score 40.4; DB 4; Length 3381;
Matches 131; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

US-09-336-447A-6
; Sequence 6, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP1 AND USP2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY.024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; PRIOR FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3381
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-09-336-447A-6

Query Match
Best Local Similarity 4.1%; Score 40.4; DB 3; Length 3381;
Matches 131; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 504 CAGGGACCAACAATAGTGTGGATCAGGGAGCAACAATGTGTTCCGGGAATGACAACA 563
Db 781 CAGCGGAGAAATCAACGGTTGCGAGGGGAGGAGTAACCAAGCCACAGGCAACAACCTCAA 840

QY 564 CGTCTGTATCTGGGAATAACAACCATGTCTTGGGAGCAACAACACTGTTGTAACCTGGAA 623
Db 841 CGGTTGAGCGGCTCTTTATAACCAAGCCACAGGCAACAACCTCAACGGTTGCAAGGTGGCT 900

QY 624 GTGCAATACTGTAGTTGGTAGCAACCATGTCTGATCAGGCAACAAGCATATTGTTACTG 683
Db 901 CTCATAACCAAGCCACAGGTGAAGTTTCATTTCGAGCAGGTGTAGAGAACAAAGCCAAATG 960

QY 684 ATAACAATAATGTGTATCCGGGAACGCAATAATGTCTGGAAGCTTCCATCTAT 743
Db 961 CCACACACCGCTCGCTCTAGGTAAATAAATACACCATCGATGCGGATACTCAGTAGCCA 1020

QY 744 CAGGGAGCAACAATACCGTATCCGGGAGCAACAATCTGTAT 785
Db 1021 TCGGCTCTAATAATACCATTTGACAGTGGCAACAACAATAATGTCT 1062

RESULT 5
US-09-336-447A-6
; Sequence 6, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP1 AND USP2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY.024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3381
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-09-336-447A-6

Query Match
Best Local Similarity 4.1%; Score 40.4; DB 3; Length 3381;
Matches 131; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 504 CAGGGACCAACAATAGTGTGGATCAGGGAGCAACAATGTGTTCCGGGAATGACAACA 563
Db 781 CAGCGGAGAAATCAACGGTTGCGAGGGGAGGAGTAACCAAGCCACAGGCAACAACCTCAA 840

QY 564 CGTCTGTATCTGGGAATAACAACCATGTCTTGGGAGCAACAACACTGTTGTAACCTGGAA 623
Db 841 CGGTTGAGCGGCTCTTTATAACCAAGCCACAGGCAACAACCTCAACGGTTGCAAGGTGGCT 900

QY 624 GTGCAATACTGTAGTTGGTAGCAACCATGTCTGATCAGGCAACAAGCATATTGTTACTG 683
Db 901 CTCATAACCAAGCCACAGGTGAAGTTTCATTTCGAGCAGGTGTAGAGAACAAAGCCAAATG 960

QY 684 ATAACAATAATGTGTATCCGGGAACGCAATAATGTCTGGAAGCTTCCATCTAT 743
Db 961 CCACACACCGCTCGCTCTAGGTAAATAAATACACCATCGATGCGGATACTCAGTAGCCA 1020

QY 744 CAGGGAGCAACAATACCGTATCCGGGAGCAACAATCTGTAT 785
Db 1021 TCGGCTCTAATAATACCATTTGACAGTGGCAACAACAATAATGTCT 1062

RESULT 6
US-09-952-267B-6
; Sequence 6, Application US/09952267B
; Patent No. 6753417
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP1 AND USP2 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY.024
; CURRENT APPLICATION NUMBER: US/09/952,267B
; CURRENT FILING DATE: 2001-09-12
; PRIOR FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3381
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-09-952-267B-6

Query Match
Best Local Similarity 4.1%; Score 40.4; DB 4; Length 3381;
Matches 131; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 504 CAGGGACCAACAATAGTGTGGATCAGGGAGCAACAATGTGTTCCGGGAATGACAACA 563
Db 781 CAGCGGAGAAATCAACGGTTGCGAGGGGAGGAGTAACCAAGCCACAGGCAACAACCTCAA 840

QY 564 CGTCTGTATCTGGGAATAACAACCATGTCTTGGGAGCAACAACACTGTTGTAACCTGGAA 623
Db 841 CGGTTGAGCGGCTCTTTATAACCAAGCCACAGGCAACAACCTCAACGGTTGCAAGGTGGCT 900

QY 624 GTGCAATACTGTAGTTGGTAGCAACCATGTCTGATCAGGCAACAAGCATATTGTTACTG 683
Db 901 CTCATAACCAAGCCACAGGTGAAGTTTCATTTCGAGCAGGTGTAGAGAACAAAGCCAAATG 960

QY 684 ATAACAATAATGTGTATCCGGGAACGCAATAATGTCTGGAAGCTTCCATCTAT 743
Db 961 CCACACACCGCTCGCTCTAGGTAAATAAATACACCATCGATGCGGATACTCAGTAGCCA 1020

QY 744 CAGGGAGCAACAATACCGTATCCGGGAGCAACAATCTGTAT 785
Db 1021 TCGGCTCTAATAATACCATTTGACAGTGGCAACAACAATAATGTCT 1062

RESULT 7
US-09-249-180-1
; Sequence 1, Application US/09249180
; Patent No. 6268548
; GENERAL INFORMATION:
; APPLICANT: Elthon, Thomas E
; APPLICANT: Lund, Adrian A
; APPLICANT: Bhatramakki, Dinakar
; APPLICANT: Rhoads, David M.
; TITLE OF INVENTION: Isolation and Characterization of Heat Shock Protein
; FILE REFERENCE: UNV52819
; CURRENT APPLICATION NUMBER: US/09/249,180
; CURRENT FILING DATE: 1999-02-12
; EARLIER APPLICATION NUMBER: 60/076/014
; PRIOR FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Created in PatentIn Ver. 2.0, Edited in WordPerfect 6.1
; SEQ ID NO 1
; LENGTH: 1028
; TYPE: DNA
```

ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (79)..(735)  
FEATURE:  
NAME/KEY: gene  
LOCATION: (1)..(1028)  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (424)..(693)  
OTHER INFORMATION: Heat Shock Domain  
FEATURE:  
NAME/KEY: exon  
LOCATION: (1)..(328)  
FEATURE:  
NAME/KEY: exon  
LOCATION: (329)..(1028)  
FEATURE:  
NAME/KEY: Poly A site  
LOCATION: (1028)..(1028)  
FEATURE:  
NAME/KEY: mRNA  
LOCATION: (1)..(1028)  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(1028)  
OTHER INFORMATION: Zea Mays L., Line B73  
FEATURE:  
NAME/KEY: transit\_peptide  
LOCATION: (79)..(213)  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: (736)..(1028)  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: (1)..(78)  
PUBLICATION INFORMATION:  
AUTHORS: Lund, Adrian A.  
AUTHORS: Blum, Paul H.  
AUTHORS: Bhattachakki, Dinakar  
AUTHORS: Elthon, Thomas E.  
TITLE: Heat-Stress Response of Maize Mitochondria  
JOURNAL: Plant Physiol.  
VOLUME: 116  
PAGES: 1097-1110  
DATE: 1998-03-00  
US-09-249-180-1

Query Match 4.0%; Score 38.8; DB 3; Length 1028;  
Best Local Similarity 52.5%; Pred. No. 0.083;  
Matches 85; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
QY 175 AGGAGCCGTCCTCCCGCGGCGATGTCGGGCGCTCATGTCGAGCTGGGAAGTGT 234  
DB 189 AGTAGCCGGCGCTCCCGGCTACACACCGGCGCTCCGCTCCGACGCTACGAGGGGC 248  
QY 235 GGGATGCCAAACAGCAAGCGCGGCTCGTGGCGTTCGCCAGTCCCAAGCGCGCTTGG 294  
DB 249 CGAGTCGAACAGCATAGCTCCGCGAGTAGCATGGGCGGACGCGCGCGGACTACGC 308  
QY 295 AGGATCATCCATGTCGATGGTGGATTCATCACTTCG 336  
DB 309 TGTGCCAGCCTGTTCTCAGATATTTTCCGTGATCCGCTTAG 350

RESULT 8  
US-09-434-288-12/c  
Sequence 12, Application US/09434288  
Patent No. 6303767  
GENERAL INFORMATION:  
APPLICANT: Betlach C., Melanie  
APPLICANT: McDaniel, Robert  
TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA

TITLE OF INVENTION: CONSTRUCTS THEREFOR  
FILE REFERENCE: 30062-20030.00  
CURRENT APPLICATION NUMBER: US/09/434,288  
CURRENT FILING DATE: 1999-11-05  
PRIOR APPLICATION NUMBER: 60/107,093  
PRIOR FILING DATE: 1998-11-05  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 12  
LENGTH: 1476  
TYPE: DNA  
ORGANISM: Streptomyces narbonensis  
US-09-434-288-12

Query Match 4.0%; Score 38.8; DB 3; Length 1476;  
Best Local Similarity 50.0%; Pred. No. 0.11;  
Matches 97; Conservative 0; Mismatches 97; Indels 0; Gaps 0;  
QY 85 CATCTTGCAGTGCAGGAGCAACGTCGTGGTGTGCCACCAAGACGACCTCCACGGTT 144  
DB 1281 CCTGCCCGCGTGAAGCGGACCCCGGCTCGGTTCTCCACGGGACATCCGCGACG 1222  
QY 145 GAGGGGCTCGCTGAGAACCTAAGCGCAAGGAGCGCTCGGCTCCGCGCGCATGTC 204  
DB 1221 CGACCTCTCGCCCGGGAATGCGGCGCTCGACGCGCTGTCTCACTTCCGCGCGAG 1162  
QY 205 CGGCGCTCATGTCGAGCTGGGAAGTGTGGATGCGAAACAGCAAGCGCGCGCTCGT 264  
DB 1161 CCAGTCGACCGCTCCATCGCGGGGCGTCCGTTTCACCGGACCAACGTCACGAGGCAC 1102  
QY 265 GGGTTCGGGCTCC 278  
DB 1101 GCAGACCTCGTCC 1088

RESULT 9  
US-10-101-464A-840  
Sequence 840, Application US/10101464A  
Patent No. 8768041  
GENERAL INFORMATION:  
APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, Nicolaas  
APPLICANT: Higgins, Colleen M.  
TITLE OF INVENTION: Compositions Isolated from Plant Cells  
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
FILE REFERENCE: 11000.1020c2  
CURRENT APPLICATION NUMBER: US/10/101,464A  
CURRENT FILING DATE: 2002-03-18  
PRIOR APPLICATION NUMBER: 09/704,302  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 09/228,986  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/162,866  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: PCT/US00/00724  
PRIOR FILING DATE: 2000-01-11  
NUMBER OF SEQ ID NOS: 989  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 840  
LENGTH: 2900  
TYPE: DNA  
ORGANISM: Pinus radiata  
US-10-101-464A-840

Query Match 4.0%; Score 38.8; DB 4; Length 2900;  
Best Local Similarity 56.2%; Pred. No. 0.16;  
Matches 73; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
QY 274 GCTCCCAAGCGCGCTTGGAGGATCATCCATGTCGATGGTGGATGATCACT 333  
DB 1723 GCTTCTTACACGCACTTGAAGTCTGATACAGCTTCAATCAGTGAGCTTCAAAATCT 1782  
QY 334 TCGCTATTTGGATCTCTCGGTAATTAATTGGTGGGAGGTACCAAAAAGTTTGCAGAT 393

```
Db      1783 TCAGGACATGGATTCTCTCCACAAATGTGACAGCTTGAGACTTCAAAGCTTTGATATT 1842
QY      394  ACGGCTCAAG 403
Db      1843 TCGGTTTCATG 1852

RESULT 10
US-10-101-464A-268
; Sequence 268, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 268
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-268

Query Match      3.9%; Score 38.4; DB 4; Length 1107;
Best Local Similarity 49.6%; Pred. No. 0.12;
Matches 127; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY      123 ACCACGACGACCTCCAGCGTTGAGGGGCTCTGAGAACCTTAAGCGGCAAGGAGCGG 182
Db      359 AGCTCCACGCGCTCATGGCACTCAAGCGCGCTCGACCCCTCCGCGCGGGTCTCCACCT 418
QY      183 TCCGCTCCGCGCGCATGTCTCGGCGCTTCATGCTGCAGCTGGGAAGGTGTGGGATGCG 242
Db      419 CGTGTCTCCGCGCGCATGTCTCGGCGCTTCATGCTGCAGCTGGGAAGGTGTGGGATGCG 478
QY      243 AAACAGCAAGCGCGCGCTGTGCGGCTTCCGCGCTCCGCGCGGGCTTTGGAGGGATCA 302
Db      479 A---CGAGCGCGCGCGTGTGTAACGTCTCGCTGCAGGGGAAGGCGCTGAGGGGCGGA 535
QY      303 TCCATGCTGATGTGAGCTGTGATCACTTCGCTATTGATCTCTCGGGTAATTTCAT 362
Db      536 TCCGCGCGGAGATCGGCGGCTCGGGAGCTGACCGGGCTGTACCTGCACCTTCAACGCC 595
QY      363 TGGTTGGGAGGTACC 378
Db      596 TGGCGCGGAGGTGCC 611

RESULT 11
US-10-101-464A-865
; Sequence 865, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
```

```
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 865
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-865

Query Match      3.9%; Score 38.4; DB 4; Length 2735;
Best Local Similarity 49.6%; Pred. No. 0.21;
Matches 127; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY      123 ACCACGACGACCTCCAGCGTTGAGGGGCTCTGAGAACCTTAAGCGGCAAGGAGCGG 182
Db      343 AGCTCCACGCGCTCATGGCACTCAAGCGCGCTCGACCCCTCCGCGCGGGTCTCCACCT 402
QY      183 TCCGCTCCGCGCGCATGTCTCGGCGCTTCATGCTGCAGCTGGGAAGGTGTGGGATGCG 242
Db      403 CGTGTCTCCGCGCGCATGTCTCGGCGCGGGCTCGTTCGAGGGGGTGGCGTGGCG 462
QY      243 AAACAGCAAGCGCGCGCTGTGCGGCTTCCGCGCTCCGCGCGGGCTTTGGAGGGATCA 302
Db      463 A---CGAGCGCGCGCGTGTGTAACGTCTCGCTGCAGGGGAAGGCGCTGAGGGGCGGA 519
QY      303 TCCATGCTGATGTGAGCTGTGATCACTTCGCTATTGATCTCTCGGGTAATTTCAT 362
Db      520 TCCGCGCGGAGATCGGCGGGCTCGGAGCTTACCGGGCTGTACCTGCACCTTCAACGCC 579
QY      363 TGGTTGGGAGGTACC 378
Db      580 TGGCGCGGAGGTGCC 595

RESULT 12
US-09-228-986-39
; Sequence 39, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-39

Query Match      3.9%; Score 38.2; DB 3; Length 454;
Best Local Similarity 51.5%; Pred. No. 0.077;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY      217 CTGCACTGGGAAGGTGTGGGATGCGAAACAGCAAGCGCGCGCTGTCGCGTTGGCGCT 276
Db      237 CTGCGCTGTGGAACGGAATTTCTGTGAGGCCAACACAGAACGAGTGTGTTCAATTTCT 296
QY      277 CCCCAAGCGCGCTTTGGAGGGATCATCCATGCTGATGTTGGAGCTTGATCACTTCG 336
```

```
Db 297 CCTGAGTCTGCTTGAATGATGCTCTCCCGTATATTGGGAACCTCTCTCTTTGGG 356
QY 337 CTATTGGATCTCTCGGTAATTCATTTGGTGGGAGGTACCAAAAAGTTT 387
Db 357 GCATTTAGATCTTCTTGAATGCTTTGAGTGGGAGAAATCCAGCAGATT 407

RESULT 13
US-10-101-464A-39
; Sequence 39, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-101-464A-39

Query Match 3.9%; Score 38.2; DB 4; Length 454;
Best Local Similarity 51.5%; Pred. No. 0.077;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 217 CTGACGTGGGAGTGTGGGATGGAACAGCAAGCGCGCTGCTGGCGTTGCGGCT 276
Db 237 CTGCTGTGGAAACGGAATTTCTGTCAGGCCCAACACGAAACGAGTGGTTTCAATTCTCT 296
QY 277 CCCCAGCGCGCTTGGAGGATCATCCCATGTCGATTTGGTGGTGGTGGTGGTGGTGGTGG 336
Db 297 CCCTGAGTGTGTTGAATGATGCTCTCCCGTATATTGGGAACCTCTCTCTTTGGCG 356
QY 337 CTATTGGATCTCTCGGTAATTCATTTGGTGGGAGGTACCAAAAAGTTT 387
Db 357 GCATTTAGATCTTCTTGAATGCTTTGAGTGGGAGAAATCCAGCAGATT 407

RESULT 14
US-10-101-464A-421
; Sequence 421, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
```

```
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 421
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-101-464A-421

Query Match 3.9%; Score 38.2; DB 4; Length 498;
Best Local Similarity 51.5%; Pred. No. 0.081;
Matches 88; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 217 CTGACGTGGGAGTGTGGGATGGAACAGCAAGCGCGCTGCTGGCGTTGCGGCT 276
Db 237 CTGCTGTGGAAACGGAATTTCTGTCAGGCCCAACACGAAACGAGTGGTTTCAATTCTCT 296
QY 277 CCCCAGCGCGCTTGGAGGATCATCCCATGTCGATTTGGTGGTGGTGGTGGTGGTGGTGG 336
Db 297 CCCTGAGTGTGTTGAATGATGCTCTCCCGTATATTGGGAACCTCTCTCTTTGGCG 356
QY 337 CTATTGGATCTCTCGGTAATTCATTTGGTGGGAGGTACCAAAAAGTTT 387
Db 357 GCATTTAGATCTTCTTGAATGCTTTGAGTGGGAGAAATCCAGCAGATT 407

RESULT 15
US-09-621-976-15639
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match 3.9%; Score 38.2; DB 4; Length 505;
Best Local Similarity 12.3%; Pred. No. 0.082;
Matches 54; Conservative 188; Mismatches 196; Indels 1; Gaps 1;

QY 184 CCGCTCCGCGCGCATGCTCGCGCTTCCCAAGCGCGCTTGGAGGATCAT 243
Db 6 SSRYRTSSKSKRYGGKYSYSSRWSMKYAWGRKYGTSGRCGSGRGMWCSKWGRYSY 65
QY 244 AACAGCAAGCGCGCTGCTGGCTTCCCAAGCGCGCTTGGAGGATCAT 303
Db 66 WGYKWSKSKMKYSGKMGTSKS-TRKYTYTSKSKRTCKYRWSWKRWRKMY 124
QY 304 CCATCGTGAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 363
Db 125 YRMYKVCYASYSYVRRCKRYTGMTRGMYGCKKRMCKSKSTRYMYRYWYKMTGACYGS 184
QY 364 GGTGGGAGGTACCAAAAAGTTTGGAGATACGGCTCAAGAGCTCACCACTGACAGCCA 423
Db 185 KGMSCGSRSKYSGWYVYGYKTYMCTSKYKSMYSYKSSMCTYMYMCTY 244
QY 424 GTCACTCGGTATGGTTCATTAACATGCTATTGCTGAGCAGTAGAAGACGCTCGA 483
Db 245 KSYVYCRKSYTYRSTSKKMGTKRWSYTWMSKSYTWGCSKKRMYWSAGAYAM 304
QY 484 TGAAGAACCAATACATATACAGGACCAACATAGTGTGGATCAGGGAGCAACATGT 543
Db 305 MSWVCARMCMAMRMSAWWKCSRAKKYMYMAKSCMYCAKMSCSARSACGRSCCTTTKY 364
```

|    |     |   |     |
|----|-----|---|-----|
| Qy | 544 | TGTTTCGGGAATGACACACGGTCGTATCTGGGAATAACACCATGTGTCTGGGAGCAA | 603 |
| Db | 365 | YKMTTKCYCWKYRCMWSMCAVCYTGWICYTKTWGTCWTKAACTVGTYSCTKMW     | 424 |
| Qy | 604 | CAACACTGTGTAACTGGA  | 622 |
| Db | 425 | YMWTTCTCWTWYGTTRA   | 443 |

Search completed: May 11, 2005, 02:55:00  
Job time : 221 secs

*This Page Blank (uspro)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 21:12:48 ; Search time 650 Seconds  
(without alignments)  
8925.148 Million cell updates/sec

Title: US-10-657-852A-3

Perfect score: 980

Sequence: 1 gcttgatccatcaaggt.....ccctttacataaaaaaaaaa 980

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001s:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002s:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003s:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004s:\*
- 13: Geneseqn2004bs:\*

\*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 980   | 100.0       | 980    | 12 | Adm41460 Perennial |
| 2          | 904.4 | 92.3        | 996    | 12 | Adm41461 Tall fesc |
| 3          | 541.6 | 55.3        | 1006   | 12 | Adm41468 Tall fesc |
| 4          | 503.6 | 51.4        | 1008   | 12 | Adm41467 Perennial |
| 5          | 492.8 | 50.3        | 1007   | 12 | Adm41469 Perennial |
| 6          | 451.2 | 46.0        | 1230   | 12 | Adm41463 Tall fesc |
| 7          | 443.6 | 45.3        | 1212   | 12 | Adm41464 Perennial |
| 8          | 421.8 | 43.0        | 1246   | 12 | Adm41466 Tall fesc |
| 9          | 419   | 42.8        | 1083   | 12 | Adm41470 Perennial |
| 10         | 417.4 | 42.6        | 1084   | 12 | Adm41462 Perennial |
| 11         | 394.4 | 40.2        | 1064   | 12 | Adm41465 Perennial |
| 12         | 270.4 | 27.6        | 959    | 12 | Adm41459 Tall fesc |
| 13         | 264.2 | 27.0        | 357    | 2  | Aax99717 Grass ant |
| 14         | 262.2 | 26.8        | 841    | 12 | Adm41458 Perennial |
| 15         | 79.4  | 8.1         | 4536   | 12 | Adj40496 Plant cDN |
| 16         | 54    | 5.5         | 2000   | 8  | Ada71938 Rice gene |
| 17         | 51.8  | 5.3         | 1138   | 1  | Aan91630 3' sequen |
| 18         | 50.6  | 5.2         | 1912   | 12 | Adj39568 Plant cDN |
| 19         | 48.2  | 4.9         | 3000   | 12 | Adi32615 Rice tran |
| 20         | 44.8  | 4.6         | 3453   | 3  | AAA79709 Eucalyptu |

|    |      |     |        |    |          |                    |
|----|------|-----|--------|----|----------|--------------------|
| 21 | 43.8 | 4.5 | 235070 | 11 | ACM45174 | Acn45174 Human gen |
| 22 | 43   | 4.4 | 7722   | 3  | AAA70168 | Aaa70168 Plasmodiu |
| 23 | 41.6 | 4.2 | 2274   | 4  | ABL17081 | Ab117081 Drosophil |
| 24 | 41.6 | 4.2 | 4274   | 4  | ABL17080 | Ab117080 Drosophil |
| 25 | 41.4 | 4.2 | 966    | 6  | ABQ17272 | Abq17272 Oligonuc1 |
| 26 | 41.4 | 4.2 | 966    | 6  | ABQ17273 | Abq17273 Oligonuc1 |
| 27 | 41   | 4.2 | 408    | 3  | AAA79411 | Aaa79411 Eucalyptu |
| 28 | 40.8 | 4.2 | 702    | 5  | AAS67541 | Aas67541 DNA encod |
| 29 | 40.4 | 4.1 | 396    | 3  | AAA79420 | Aaa79420 Eucalyptu |
| 30 | 40.4 | 4.1 | 3381   | 2  | AAV41343 | Aav41343 M. catarr |
| 31 | 40.4 | 4.1 | 12614  | 4  | AAI99322 | Aai99322 Human exc |
| 32 | 40.4 | 4.1 | 12614  | 5  | AAI63672 | Aai63672 Human kid |
| 33 | 40.2 | 4.1 | 3228   | 5  | AAS81183 | Aas81183 DNA encod |
| 34 | 40   | 4.1 | 804    | 12 | ADM80069 | Adm80069 Spiramyci |
| 35 | 40   | 4.1 | 804    | 12 | ADN97585 | Adn97585 S ambofac |
| 36 | 40   | 4.1 | 30943  | 12 | ADM80034 | Adm80034 Spiramyci |
| 37 | 40   | 4.1 | 30943  | 12 | ADN97550 | Adn97550 S ambofac |
| 38 | 39.6 | 4.0 | 2000   | 8  | ADA71938 | Ada71938 Rice gene |
| 39 | 39.6 | 4.0 | 2298   | 4  | ABL18481 | Ab118481 Drosophil |
| 40 | 39.6 | 4.0 | 2479   | 11 | ADM03281 | Adm03281 Human cDN |
| 41 | 39.6 | 4.0 | 4298   | 4  | ABL18480 | Ab118480 Drosophil |
| 42 | 39.2 | 4.0 | 1443   | 10 | ABZ23668 | Abz23668 H. pylori |
| 43 | 39.2 | 4.0 | 2274   | 4  | ABL17081 | Ab117081 Drosophil |
| 44 | 39.2 | 4.0 | 4274   | 4  | ABL17080 | Ab117080 Drosophil |
| 45 | 38.8 | 4.0 | 1028   | 3  | AAA50251 | Aaa50251 Maize hea |

## ALIGNMENTS

RESULT 1  
ADM41460  
ID ADM41460 standard; cDNA; 980 BP.  
XX  
AC ADM41460;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Perennial ryegrass antifreeze protein AFP2 cDNA.  
XX  
KW Antifreeze; perennial ryegrass; cold tolerance; plant;  
KW antigout; litholytic; nephrotropic; cytosstatic; gene; ss.  
XX  
OS Lolium perenne.  
FH Key Location/Qualifiers  
FT CDS 35..838  
FT /\*tag= b  
FT /\*product= "Antifreeze protein AFP2"  
FT sig\_peptide 35..103  
FT /\*tag= a  
FT mat\_peptide 104..835  
FT /\*tag= c

WO2004022700-A2.

18-MAR-2004.

09-SEP-2003; 2003WO-NZ000199.

09-SEP-2002; 2002US-0409557P.

(GENE-) GENESIS RES & DEV CORP LTD.

(WRIG-) WRIGHTSON SEEDS LTD.

Demmer J, Shenk MA, Hall C, Fish SA;

WPI: 2004-248453/23.

P-PSDB; ADM41473.

XX New antifreeze proteins and encoding polynucleotides, useful for  
PT modulating cold tolerance in organisms, as food additives, or for  
PT treating tumors or disorders associated with the presence of unwanted

PT biocrystals (e.g. gout).  
 XX Claim 1; SEQ ID NO 3; 71pp; English.  
 XX The present sequence is that of cDNA encoding APP2, an antifreeze protein  
 CC of perennial ryegrass. The cDNA was isolated from a leaf blade cDNA  
 CC expression library. The invention provides forage grass (perennial  
 CC ryegrass and tall fescue) antifreeze proteins and the polynucleotides  
 CC encoding them ADM41458-ADM41463. The polynucleotides were isolated from  
 CC tissues taken at different times of year (winter and spring) and from  
 CC different parts of the plants. The polynucleotides can be used to  
 CC modulate the cold tolerance of an organism, especially plants, mammals,  
 CC insects, fungi, archaea and bacteria. The method involves incorporating  
 CC an antifreeze polynucleotide, under the control of a gene promoter  
 CC sequence, into the genome of the organism, or introducing double-stranded  
 CC RNA corresponding to the polynucleotide into the cells of the organism,  
 CC thereby inhibiting expression of an antifreeze polypeptide. The  
 CC antifreeze protein can be used for the cryopreservation of a cell or  
 CC tissue, as a food additive of a frozen food product, in a method for  
 CC decreasing the time required to dehydrate a composition, to treat a  
 CC disorder characterised by biocrystals associated with disorders such as  
 CC gout and kidney stones, to preserve the viability of a molecular biology  
 CC reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and  
 CC to protect a plant from damage due to frost or freezing.  
 XX SQ Sequence 980 BP; 264 A; 223 C; 253 G; 240 T; 0 U; 0 Other;

Query Match 100.0%; Score 980; DB 12; Length 980;  
 Best Local Similarity 100.0%; Pred. No. 9e-297;  
 Matches 980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTTCATTCCTCAATCAAGGTTCTTGTTCATCCATGCTGCAATACATGGCAAGTGTG 60  
 DB 1 GCTTCATTCCTCAATCAAGGTTCTTGTTCATCCATGCTGCAATACATGGCAAGTGTG 60  
 QY 61 CATGCTGCTGTTCTTGGGTTTCATCTTCAGGTGGCAGGCAACGCTGCTGCTG 120  
 DB 61 CATGCTGCTGTTCTTGGGTTTCATCTTCAGGTGGCAGGCAACGCTGCTGCTG 120  
 QY 121 CCACACAGACCTCCACGGTTGAGGGGCTGCTGAGAACCTTAAGCGGCAAGGAGC 180  
 DB 121 CCACACAGACCTCCACGGTTGAGGGGCTGCTGAGAACCTTAAGCGGCAAGGAGC 180  
 QY 181 CGTCCGCTCCGCGCGCATGCTCGCGCGCTCATGCTGCAAGTGGGAGGTGGGATG 240  
 DB 181 CGTCCGCTCCGCGCGCATGCTCGCGCGCTCATGCTGCAAGTGGGAGGTGGGATG 240  
 QY 241 CGAAACAGCAAGCGCGCGCTGCTGGCGTTGCGGCTCCCAAGCGCGGCTTGGAGGAT 300  
 DB 241 CGAAACAGCAAGCGCGCGCTGCTGGCGTTGCGGCTCCCAAGCGCGGCTTGGAGGAT 300  
 QY 301 CATCCATGCTGATGTGAGCTTGATCAGCTTGGCTATTTGGATCTCTCGGTAATTC 360  
 DB 301 CATCCATGCTGATGTGAGCTTGATCAGCTTGGCTATTTGGATCTCTCGGTAATTC 360  
 QY 361 ATTGTTGGGAGGTACCAAAAGTTTCAGATACGGCTCAAGAGCCTCACCCTGACAG 420  
 DB 361 ATTGTTGGGAGGTACCAAAAGTTTCAGATACGGCTCAAGAGCCTCACCCTGACAG 420  
 QY 421 CCAGTCACTCGGTATGGTTCCATTAAATGCTATTTGATGTCAGTGAAGAACGCT 480  
 DB 421 CCAGTCACTCGGTATGGTTCCATTAAATGCTATTTGATGTCAGTGAAGAACGCT 480  
 QY 481 CGATGAGAACCAATACAAATATCAGGACCAACAAATAGTTGATCAGGAGCAACAA 540  
 DB 481 CGATGAGAACCAATACAAATATCAGGACCAACAAATAGTTGATCAGGAGCAACAA 540  
 QY 541 TGTGTTTCCGGGATGACACACGCTGCTATCTGGGAATAACAACTGCTCTGGGAG 600  
 DB 541 TGTGTTTCCGGGATGACACACGCTGCTATCTGGGAATAACAACTGCTCTGGGAG 600  
 QY 601 CAACAACACTGTTGTAATGGAAGTGACAATCTGATGTTGAGTACCAACCAATGCTGATC 660  
 DB 601 CAACAACACTGTTGTAATGGAAGTGACAATCTGATGTTGAGTACCAACCAATGCTGATC 660

DB 601 CAACAACACTGTTGTAATGGAAGTGACAATCTGATGTTGAGTACCAACCAATGCTGATC 660  
 QY 661 AGGACAAAGCATATTTGTTACTGATAACAATAATGTTGTTATCCGGGAACGACAATAATGT 720  
 DB 661 AGGACAAAGCATATTTGTTACTGATAACAATAATGTTGTTATCCGGGAACGACAATAATGT 720  
 QY 721 GTCTGGAAGCTTCCTCACTACTGTTATCAGGGGACCAATAACCGTATCCGGGACCAATAAC 780  
 DB 721 GTCTGGAAGCTTCCTCACTACTGTTATCAGGGGACCAATAACCGTATCCGGGACCAATAAC 780  
 QY 781 TGTATCCGGGAGCAACCATATCGTATCTGCGGAGCAACAAAGTTCGTAACAGATGGTTAATA 840  
 DB 781 TGTATCCGGGAGCAACCATATCGTATCTGCGGAGCAACAAAGTTCGTAACAGATGGTTAATA 840  
 QY 841 TTCTGTAGTGCAGGATGCTTCCATCTTCCCAAGTTCAGTGTAGCTTACAAATAGA 900  
 DB 841 TTCTGTAGTGCAGGATGCTTCCATCTTCCCAAGTTCAGTGTAGCTTACAAATAGA 900  
 QY 901 TGGAGACAATCACGTTATGTAACCTTCAGGATATGCAATCTTTCCTTTAATAAAGCTT 960  
 DB 901 TGGAGACAATCACGTTATGTAACCTTCAGGATATGCAATCTTTCCTTTAATAAAGCTT 960  
 QY 961 CCCTTTACATAAAAAAAA 980  
 DB 961 CCCTTTACATAAAAAAAA 980

RESULT 2  
 ADM41461  
 ID ADM41461 standard; cDNA; 996 BP.  
 XX  
 AC ADM41461;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Tall fescue antifreeze protein cDNA.  
 XX  
 KW Antifreeze; fescue; cold tolerance; transgenic; plant; antigout;  
 KW litholytic; nephrotropic; cyostatic; gene; ss.  
 XX  
 OS Schedonorus arundinaceus.  
 XX  
 FH Key Location/Qualifiers  
 CDS 41..850  
 FT /\*tag= b  
 FT /product= "Antifreeze protein"  
 FT /transl\_except= (pos:410..412,aa:Xaa)  
 FT /note= "Xaa= unknown"  
 FT sig\_peptide 41..112  
 FT /\*tag= a  
 FT mat\_peptide 113..847  
 FT /\*tag= c  
 XX  
 FN WO2004022700-A2.  
 XX  
 PD 18-MAR-2004.  
 XX  
 PF 09-SEP-2003; 2003WO-NZ000199.  
 XX  
 PR 09-SEP-2002; 2002US-0409557P.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (WRIG-) WRIGHTSON SEEDS LTD.  
 XX  
 PI Demmer J, Shenk MA, Hall C, Fish SA;  
 XX P-PSDB; ADM41474.  
 DR WPI; 2004-248453/23.  
 XX  
 PT New antifreeze proteins and encoding polynucleotides, useful for  
 PT modulating cold tolerance in organisms, as food additives, or for  
 PT treating tumors or disorders associated with the presence of unwanted  
 PT biocrystals (e.g. gout).

| XX | Claim 1; SEQ ID NO 4; 71pp; English.                                       |
|----|--|
| PS | The present sequence is that of cDNA encoding an antifreeze protein of     |
| CC | tall fescue. The cDNA was isolated from a leaf blade cDNA expression       |
| CC | library. The invention provides forage grass (perennial ryegrass and tall  |
| CC | fescue) antifreeze proteins and the polynucleotides encoding them          |
| CC | CC ADM41458-ADMA1483. The polynucleotides were isolated from tissues taken |
| CC | at different times of year (winter and spring) and from different parts    |
| CC | of the plants. The polynucleotides can be used to modulate the cold        |
| CC | tolerance of an organism, especially plants, mammals, insects, fungi,      |
| CC | archaea and bacteria. The method involves incorporating an antifreeze      |
| CC | polynucleotide, under the control of a gene promoter sequence, into the    |
| CC | genome of the organism, or introducing double-stranded RNA corresponding   |
| CC | to the polynucleotide into the cells of the organism, thereby inhibiting   |
| CC | expression of an antifreeze polypeptide. The antifreeze protein can be     |
| CC | used for the cryopreservation of a cell or tissue, as a food additive or   |
| CC | a frozen food product, in a method for decreasing the time required to     |
| CC | dehydrate a composition, to treat a disorder characterised by biocrystals  |
| CC | associated with disorders such as gout and kidney stones, to preserve the  |
| CC | viability of a molecular biology reagent, to destroy unwanted tissue in a  |
| CC | patient e.g. tumour tissue, and to protect a plant from damage due to      |
| CC | frost or freezing.   |
| XX |  |
| SQ | Sequence 996 BP; 268 A; 235 C; 250 G; 243 T; 0 U; 0 Other;                 |
|    | Query Match 92.3%; Score 904.4; DB 12; Length 996;                         |
|    | Best Local Similarity 96.2%; Pred. No. 5.2e-273;                           |
|    | Matches 949; Conservative 0; Mismatches 31; Indels 6; Gaps 2               |
| Qy | 1 GCTTGCATTCGAATCAAGGTTCTTGTTCATTCATGCTGTAATACATGCGAAAGTGTG 60             |
| Db | 7 GCTTGCATTCGAATCAAGGTTCTTGTTCATTCATGCTGTAATACATGCGAAAGTGTG 66             |
| Qy | 61 CATGCTGTGTCTTCTTTGGGGTTCA---TCTTGCAGGTGGCAGCAGCAACTCTGTGGTTC 117        |
| Db | 67 CATGCTGTGTCTTCTTTGGGCTTCATCTCTTGCAGGTGGCGGAGCAACGTCTGTGGTC 126          |
| Qy | 118 GTGCCACACGACGACCTCCACGCTTGAGGGGCTCGTGTAGAACTTAAGCGGCAAAAGG 177         |
| Db | 127 GTGCCACACGACGACCTCCGGCATTGAGGGGCTTCGCCGAGAACCTTAAGCGGCAAAAGG 186       |
| Qy | 178 AGCGGTCCGCTCCGGCGCGAATGTCGGGCGCTCATGCTGCAGCTGGGAAGTGTGGG 237           |
| Db | 187 AGCGGTCCGCTCCGGCGCGAATGTCGGGCGCTCATGCTGCAGCTGGGAAGTGTGGG 246           |
| Qy | 238 ATCGGAACAGCAGCGCGGCTCGTGGCGTTCGGGCTCCCAAGCGCGGCTTGGAGG 297             |
| Db | 247 ATCGGAACAGCAGCGCGGCTCGGCGGCTTCGGGCTCCCAAGCGCGGCTTGGAGG 306             |
| Qy | 298 GATCATCCCATCGTCGATTGGTGAGCTTGATCACTTCGCTATTTTGGATCTCTCGGGTAA 357       |
| Db | 307 GACCATCCCATCGTCGATTGGTGAGCTTGATCACTTCGTTGTTGGATCTCTCGGGTAA 366         |
| Qy | 358 TTCAATGTTGGGAGGTAACAAAAGTTTGAGATACGGCTCAAGAGCCTCACCACATGA 417          |
| Db | 367 TTCAATGTTGGGAGGTAACAAAAGTTTGAGATACGGCTCAAGAGCCTCACCACATGA 426          |
| Qy | 418 CAGCCAGTCACTCGGTATGGGTTCCATTAAACATGCTATTGATGTAGCAGT---AGAG 474         |
| Db | 427 TGGCCAGTCACTCGGTATGGGTTCCATTAAACATGCTATTGATGTAGCAGTAAACAGAG 486        |
| Qy | 475 AACGCTCGATGAAGAACCAATACAAATATCAGGGACCAACATAGTGTGGATCAGGGAG 534         |
| Db | 487 AACCTTCGATGAAGAACCAATACAAATATCAGGGACCAACATAGTGTGGATCAGGGAG 546         |
| Qy | 535 CAACAATGTTGTTCCGGGAATGACAACAACGGTTCGATCTCGGGAATAAACAACCATGTGTC 594     |
| Db | 547 CAACAATGTTGTTCCGGGAATGACAACAACGGTTCATATCTGGGAATAAACAACCATGTGTC 606     |
| Qy | 595 TGGGAGCAACAACTGTTGTAACTGGGAAGTGACAAATCTGTAGTTGGTAGCAACCATGTG 654       |
| Db | 607 TGGGAGCAACAACTGTTGTAACTGGGAAGTGACAACTTTAGTTGGTAGCAACCATGTG 666         |

CC The present sequence is that of cDNA encoding an antifreeze protein of  
 CC tall fescue. The cDNA was isolated from a leaf blade cDNA expression  
 CC library. The invention provides forage grass (perennial ryegrass and tall  
 CC fescue) antifreeze proteins and the polynucleotides encoding them  
 CC ADM1448-ADM1493. The polynucleotides were isolated from tissues taken  
 CC at different times of year (winter and spring) and from different parts  
 CC of the plants. The polynucleotides can be used to modulate the cold  
 CC tolerance of an organism, especially plants, mammals, insects, fungi,  
 CC archaea and bacteria. The method involves incorporating an antifreeze  
 CC polynucleotide, under the control of a gene promoter sequence, into the  
 CC genome of the organism, or introducing double-stranded RNA corresponding  
 CC to the polynucleotide into the cells of the organism, thereby inhibiting  
 CC expression of an antifreeze polypeptide. The antifreeze protein can be  
 CC used for the cryopreservation of a cell or tissue, as a food additive of  
 CC a frozen food product, in a method for decreasing the time required to  
 CC dehydrate a composition, to treat a disorder characterised by biocrystals  
 CC associated with disorders such as gout and kidney stones, to preserve the  
 CC viability of a molecular biology reagent, to destroy unwanted tissue in a  
 CC patient e.g. tumour tissue, and to protect a plant from damage due to  
 CC frost or freezing.

SQ Sequence 1006 BP; 248 A; 261 C; 260 G; 237 T; 0 U; 0 Other;

Query Match 55.3%; Score 541.6; DB 12; Length 1006;

Best Local Similarity 76.5%; Pred. No. 5.3e-159;

Matches 756; Conservative 0; Mismatches 194; Indels 38; Gaps 6;

QY 22 TCTTGTCAATCATCGCTGATACATCGCAAGTTGTCATGCTGCTGCTCTCTGGG 81  
 DB 28 TCTATAGCATAGCATGCTGATATCCATGCGGAATGCTTGATGCTGCTCTCTCGC 87  
 QY 82 GTTCATCTTCAGAGTGGCAGGA---GCAACGTCGTGTCGTGCCACACGACGACTCCA 138  
 DB 88 GTTCTCTTTCGGGGCTTGGCAGCGCGGCGGCGCATGCCACCGGGATGACTTCG 147  
 QY 139 CCGCTTGAGGGCTCTGCTGAGAACTTAAGCGGCAAGAGCGCTCCGCCCTCCGCGCGC 198  
 DB 148 TCGCTGCGGGGCTTCGCTGAGAACTCGGGCGCGCGGCGCACTCAGCCTTCGCGCGC 207  
 QY 199 ATGTCGCGCGCTCATCTGCTGAGTGGAGGTTGGGATGCGAAGCAAGCGGCGG 258  
 DB 208 GTGGTCAGCGCGCTCATGCTGATTTGGAGAGCGCTTGGCTGCGACGGTGGCGCGCG 267  
 QY 259 CQCTGGCGTTGCGGCTCCCAAGCGCGCTTCGAGGGATCATCCCATCTGTCGATTGG 318  
 DB 268 TGTACGGCTTGTGGCTCCCGAGGCGGCTCAGCGGGCCATCCCATCTGTTGATTG 327  
 QY 319 TGAGCTTGATCACCTTCGCTATTGGATCTCTCGGGTAATTCATTGGTTGGGAGGTACC 378  
 DB 328 TCAGCTTCACCACTTACGCTACTTGGATCTTTTTCAGGTAATGCAITGGTTGGCGAGGTACC 387  
 QY 379 AAAAAGTTTGAGATACGCGCTCAGAGCGCTCACCACCTCAGAGCCAGTCACTCGGTATGGG 438  
 DB 388 CAAGAACTTCGAGGTACAGCTCAAGAGGCTCACCGCTCCCGGTC-----GTTGGG 438  
 QY 439 TTCCATTAACTGCTATTGCTGTA---GCAGTAGAAGAACGCTCGGATGAAGAACCAA 495  
 DB 439 TTTCCACCAATGCAATGTCATGATGCTGTAACAGAGATCATCTCGACGAGCGCCAA 498  
 QY 496 TACAATATCAGGACCAACAATAGTTGGATCAGGAGCAACAATGTTGTTTCGGGAA 555  
 DB 499 TACAATATCTGGAGCAACAATCTGTGATGATCGGGAGCAAAAATGTTGTTGCTGGAA 558  
 QY 556 TGACAACACGTCGCTATCTGGGATAACAACCTGTCGTGGAGCAACACACTGTTGT 615  
 DB 559 TGAACAACCGCTATCTGCGGACACAATAGTGTGTCGGAGCAACAACACTGTCGT 618  
 QY 616 AACTGGAAGTGACATACTGTAGTTGGTAGCAACCATGTCGTATCAGGAGCAAAAGCATAT 675  
 DB 619 AAGTGGAGTGACATACTCGTAACCTGGCAGCAACCATGTCGTATCAGGAGCAACACATAT 678  
 QY 676 TGTTACTGATTAACAATATGTTGTTATCCGGGAAACGACAATATGTTGTTGGAAGTTCGA 735  
 DB 735

DB 679 CGTTACAGACAAACAATAACGTTATCCGGGAACGATAAATATGTTATCCGGAGCTTTCA 738  
 QY 736 TACTGTATCAGGGAGGACATACCGTATCCGGAGGCAACAATACTGTATCCGGAGCAA 795  
 DB 739 TACCGTATCCGGGGGACAAATACCGTCTCTGGGAGCAACAATACCGTATCTGGGAGCAA 798  
 QY 796 CCATATCGTATCTCGGAGCAACAAGTCTGAACAGATGGTTAATATTTCTGTAGGTGCAAG 855  
 DB 799 CCAATGTTGATCTCGAAGCAACAAGTCTGACAGACGCTTAATGATCTGTGCGGCAATG 858  
 QY 856 ATTGCTTCCACTTT-----CCCAAGTTCAGTGTAGCTTCAATCAA 896  
 DB 859 ATTGTTTCCACTTAACTAGCTCAGCTTCGTTCTTGTCCAAAGTTCACTGTACCTCAGAGTCAG 918  
 QY 897 TAGATGGAGCAATCAGTTATGTAACCTCA-GGATATGGCATACTTTTC---CTTTAA 952  
 DB 919 TTGGTGGCTTCAATCGGGGTATGTAACCTTCAATGATATACCATACTTTTCTCTACTATATA 978  
 QY 953 TAAAGCTTCCCTTTTACATAAAAAAAA 980  
 DB 979 TAAATTTCCCTTTAAAAAAAANA 1006

RESULT 4

ADM14467

ID ADM14467 standard; cDNA; 1008 BP.

XX ADM14467;

AC ADM14467;

DT 03-JUN-2004 (first entry)

XX Perennial ryegrass antifreeze protein APP5 cDNA.

XX Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;

XX antigout; litholytic; nephrotropic; cytostatic; gene; ss.

XX Lolium perenne.

XX Key Location/Qualifiers

XX CDS 79..843

XX /\*tag= b

XX /product= "Antifreeze protein APP5"

XX sig\_peptide 79..147

XX /\*tag= a

XX mat\_peptide 148..840

XX /\*tag= c

XX WO2004022700-A2.

XX 18-MAR-2004.

XX 09-SEP-2003; 2003WO-NZ000199.

XX 09-SEP-2002; 2002US-0409557P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX (WRIGHTSON SEEDS LTD.

XX Demmer J, Shenk MA, Hall C, Fish SA;

XX WPI; 2004-248453/23.

XX P-PSDB; ADM14480.

XX New antifreeze proteins and encoding polynucleotides, useful for  
 XX modulating cold tolerance in organisms, as food additives, or for  
 XX treating tumors or disorders associated with the presence of unwanted  
 XX biocrystals (e.g. gout).

XX Claim 1; SEQ ID NO 10; 71pp; English.

XX The present sequence is that of cDNA encoding APP5, an antifreeze protein  
 XX of perennial ryegrass. The cDNA was isolated from a leaf blade cDNA  
 XX expression library. The invention provides forage grass (perennial

|          |   |
|----------|---|
| RESULT 5 |   |
| ADM41469 |   |
| ID       | ADM41469 standard; cDNA; 1007 BP.   |
| XX       |   |
| AC       | ADM41469;   |
| XX       |   |
| DT       | 03-JUN-2004 (first entry)   |
| XX       |   |
| DE       | Perennial ryegrass antifreeze protein cDNA.                               |
| XX       |   |
| KW       | Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;        |
| KW       | antigout; litholytic; nephrotropic; cytostatic; gene; ss.                 |
| XX       |   |
| OS       | Lolium perenne.   |
| XX       |   |
| FH       | Key Location/Qualifiers   |
| FT       | CDS 72..842   |
| FT       | /*tag= b  |
| FT       | /product= "Antifreeze protein"  |
| FT       | sig_peptide 72..140   |
| FT       | /*tag= a  |
| FT       | mat_peptide 141..839  |
| FT       | /*tag= c  |
| XX       |   |
| XX       | WO2004022700-A2.  |
| PN       |   |
| XX       |   |
| PD       | 18-MAR-2004.  |
| XX       |   |
| PF       | 09-SEP-2003; 2003WO-NZ000199.   |
| XX       |   |
| PR       | 09-SEP-2002; 2002US-0409557P.   |
| XX       |   |
| PA       | (GENE-) GENESIS RES & DEV CORP LTD.                                       |
| PA       | (WRIG-) WRIGHTSON SEEDS LTD.  |
| XX       |   |
| PI       | Demmer J, Shenk MA, Hall C, Fish SA;                                      |
| XX       |   |
| DR       | WPI; 2004-248453/23.  |
| DR       | P-PSDB; ADM41482.   |
| XX       |   |
| PT       | New antifreeze proteins and encoding polynucleotides, useful for          |
| PT       | modulating cold tolerance in organisms, as food additives, or for         |
| PT       | treating tumors or disorders associated with the presence of unwanted     |
| PT       | biocrystals (e.g. gout).  |
| XX       |   |
| PS       | Claim 1; SEQ ID NO 12; 71pp; English.                                     |
| XX       |   |
| CC       | The present sequence is that of cDNA encoding, an antifreeze protein of   |
| CC       | perennial ryegrass. The cDNA was isolated from a leaf and pseudostem cDNA |
| CC       | expression library. The invention provides forage grass (perennial        |
| CC       | ryegrass and tall fescue) antifreeze proteins and the polynucleotides     |
| CC       | encoding them ADM41458-ADM41483. The polynucleotides were isolated from   |
| CC       | tissues taken at different times of year (winter and spring and from      |



CC genome of the organism, or introducing double-stranded RNA corresponding  
 CC to the polynucleotide into the cells of the organism, thereby inhibiting  
 CC expression of an antifreeze polypeptide. The antifreeze protein can be  
 CC used for the cryopreservation of a cell or tissue, as a food additive of  
 CC a frozen food product, in a method for decreasing the time required to  
 CC dehydrate a composition, to treat a disorder characterised by biocrystals  
 CC associated with disorders such as gout and kidney stones, to preserve the  
 CC viability of a molecular biology reagent, to destroy unwanted tissue in a  
 CC patient e.g. tumour tissue, and to protect a plant from damage due to  
 CC frost or freezing.  
 XX  
 SQ Sequence 1230 BP; 327 A; 279 C; 298 G; 326 T; 0 U; 0 Other;

Query Match 46.0%; Score 451.2; DB 12; Length 1230;  
 Best Local Similarity 72.0%; Pred. No. 1.5e-130;  
 Matches 708; Conservative 0; Mismatches 198; Indels 78; Gaps 6;  
 33 CCATGCTGTAATCATGCGCAAGTGTTCATGCTGTCTTCTTGGGGTTCATCTTGC 92  
 62 CCACAACTCAATCCATGCGCAATGTTGGTGTCTGTCTTCTTCTTGGTCTCTTGC 121  
 93 AGGTGCGAGGAGCAAGCTGTGTCTGTCCACACGACGACCTCCACGCGTTGAGGGCC 152  
 122 CGCCCGGAGCGGAC-----GTCTGCCACCCCGATGACCTCCGCGCGCTCGTGGCT 175  
 153 TCCTCTGAGAACCTTAAGCGGCAAGGAGCGCTCGCGCGCATGCTCCGCGCGCT 212  
 176 TTGTCGGGAACCTCAATGGC---GGGGGTCTCTTCTTCATGGAGCATGTCGGCTCT 232  
 213 CATGCTGCGAGCTGGGAGGTGTGGATGCGAAGCAGCAAGCGCGCGCTGTGGCGTTGC 272  
 233 TATGTTGCGCTCGGGAAGCGTGGGTGCGATGGAACAAGCGCGCGCTCACGCGCATTCG 292  
 273 GGCTCC-----CCAAAGCGG 287  
 293 GGCTTCGATTACCTCAAGGATGCGGCAAGCTCAAGTCTCAACCTTGCCCAACGATA 352  
 288 GCCTTCGAGGAGATCATCCCATCGTATGCTGAGCTTGATCACCCTTGTCTTTCGATC 347  
 353 GACTGCTGGCACCATCCCGTCTGATGCTGAGCTTGACCACTTTGCTACTTGTTC 412  
 348 TCTCGGGTAATTCATTTGTTGGGAGGTACCAAAAGTTTGCAGATACGGCTCAAGAGCC 407  
 413 TCTCGGATAATTCATTTGTTGTAAGGTACCCAAATAGTTTGCAGATAAGCACTCAAGGCG 472  
 408 TCACCACTGACAGCCAGTCACTCGGTATGGGTTCCATTAACATGCTTATGCAATGCA--- 464  
 473 TCGCCACCGCTGCTGCTCACTAGGTATGGCTTTCCGTAACATGCTTGCATGTGAAGG 532  
 465 GCAGTAGAAGAACGCTCGATGAAGAACCAAAATACAATATCAGGGAACAACATAGTGTG 524  
 533 GCACCCGAGAACACTTCGACGAAACAACAATATACATACAGGGGACAACACAGATTA 592  
 525 GATCAGGAGCAACAATGTTGTTTCGGGAATGACACACGCTGATCTGCGGAATAACA 584  
 593 GATCTGGGAACGCAATGCAAGTTCTTGGGAACGACAACACTGATATGTGGGAACAACA 652  
 585 ACCATGCTGTGGGAGCAACAACACTGTTGTAATCTGGAAGTGAATAGTGTGTA 644  
 653 ACATGCTGTCTGGGAGCAACAATATCTGTTATCTGGGAGTGAATATCGTAACTGGTA 712  
 645 GCAACCATGTCGTATCAGGGAACAACATATGTTACTGATAACAATATGTTGATTCG 704  
 713 GCAACCATGTTGATGTGGGACCAACAATATCATATCACTGATACAAATGACGTATCG 772  
 705 GGAACGCAACAATATGTTCTGGAAGCTTCCATCTGATCAGGGGAGCAACAATACCGTAT 764  
 773 GGAACGATAACAATATGTTCTGGGAGTTCCCATCTGATCTCCGGGAGCCAAATACTGTAT 832  
 765 CCGGAGCAACAATATCTGATCTCGGAGCAACCATATCTGATCTGGGAGCAACAAGTGC 824  
 833 CCGGAGGTAACAACATGTTATCTTGGGAACCAACATGCTGATCTGGGAGCAACAAGTGC 892

QY 825 TAACAGATGTTAATATTCTGTAGG-----TGCAAGGATTGCTTCC 864  
 DB 893 TGACAGGAGATCAATGATTTGTTCAGGGATCGTTTCCATCTTCCCTAAAGGAGCACTCAC 952  
 QY 865 ATCTTCCCAAGTTCAGTGTAGCTTCAATCAATAGATGAGACAATCAGTTATGTAAC 924  
 DB 953 CTTGTGTCAGTTCAGTGTAGCTCAGTCACTTGTGAGGACAATCGAGTTATGTAAC 1012  
 QY 925 TCA-GGATATGCAATCTTTTCT 947  
 DB 1013 TCATGATATAGCTCAATTTTCT 1036  
 RESULT 7  
 ADM41464  
 ID ADM41464 standard; cDNA; 1212 BP.  
 XX ADM41464;  
 XX DT 03-JUN-2004 (first entry)  
 XX Perennial ryegrass antifreeze protein cDNA.  
 XX DE Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;  
 XX KW antitumor; litholytic; nephrotropic; cytotatic; gene; ss.  
 XX OS Lolium perenne.  
 XX FH Key Location/Qualifiers  
 FT CDS 59..892  
 FT /\*tag= b  
 FT /\*product= "Antifreeze protein"  
 FT sig\_peptide 59..127  
 FT /\*tag= a  
 FT mat\_peptide 128..889  
 FT /\*tag= c  
 XX WO2004022700-A2.  
 XX PN 18-MAR-2004.  
 XX PD 09-SEP-2003; 2003WO-NZ000199.  
 XX PR 09-SEP-2002; 2002US-0409557P.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 XX PA (WRIG-) WRIGHTSON SEEDS LTD.  
 XX PI Demmer J, Shenk MA, Hall C, Fish SA;  
 XX WPI: 2004-248453/23.  
 XX P-PSDB; ADM41477.  
 XX New antifreeze proteins and encoding polynucleotides, useful for  
 XX modulating cold tolerance in organisms, as food additives, or for  
 XX treating tumors or disorders associated with the presence of unwanted  
 XX biocrystals (e.g. gout).  
 XX Claim 1; SEQ ID NO 7; 71pp; English.  
 XX The present sequence is that of cDNA an antifreeze protein of perennial  
 XX ryegrass. The invention provides forage grass (perennial ryegrass and  
 XX tall fescue) antifreeze proteins and the polynucleotides encoding them  
 XX ADM41458-ADM41483. The polynucleotides were isolated from tissues taken  
 XX at different times of year (winter and spring) and from different parts  
 XX of the plants. The polynucleotides can be used to modulate the cold  
 XX tolerance of an organism, especially plants, mammals, insects, fungi,  
 XX archaea and bacteria. The method involves incorporating an antifreeze  
 XX polynucleotide, under the control of a gene promoter sequence, into the  
 XX genome of the organism, or introducing double-stranded RNA corresponding  
 XX to the polynucleotide into the cells of the organism, thereby inhibiting  
 XX expression of an antifreeze polypeptide. The antifreeze protein can be  
 XX used for the cryopreservation of a cell or tissue, as a food additive of

CC a frozen food product, in a method for decreasing the time required to  
CC dehydrate a composition, to treat a disorder characterised by biocrystals  
CC associated with disorders such as gout and kidney stones, to preserve the  
CC viability of a molecular biology reagent, to destroy unwanted tissue in a  
CC patient e.g. tumour tissue, and to protect a plant from damage due to  
CC frost or freezing.  
SQ Sequence 1212 BP; 326 A; 274 C; 287 G; 325 T; 0 U; 0 Other;

Query Match 45.3%; Score 443.6; DB 12; Length 1212;  
Best Local Similarity 71.2%; Pred. No. 3.7e-128;  
Matches 706; Conservative 0; Mismatches 209; Indels 77; Gaps 6;

QY 33 CCATGCGCTGAATACATGCAAGCTTTCGATGCTGCTGCTCTCTTGGGGTTTCATCTTGC 92  
DB 45 CACACACTTATTCATGCGAATATGCTGCTGCTCTCTCTTGGGTCTCTCTTTC 104  
QY 93 AGGTGGCAGGAGCAACGTCGTGCTGCTGCCACCGACCTCCACCGCTTGAGGGGCC 152  
DB 105 TGGCGCTGTGCGGACA-----TCGTGCCACCCAGATGACCTCGCGCGCTGCGGGCT 158  
QY 153 TCGCTGAGAACTAAGCGGAAAGAGCGCTCGGCTCGCGCGCGCATGCTCGCGCGCT 212  
DB 159 TTGTGCGGAACCTCAATGGC---GGGGGTGCTCTTCTCCGTGAACATGCTGCTGCTCT 215  
QY 213 CATGCTGCAGCTGGGAAGGTGTGGGATCGGAAACAGCAAGCGCGCTGCTGCGGTTC 272  
DB 216 CATGCTGCGCTGGGAAGGTGTGGCTCGCATGTGTAACAGCGCGCGCTCACGCGGTTC 275  
QY 273 GGCTCC-----CCAGCGCG 287  
DB 276 GGCTTCGATTAGCCTCGAGGATTGCGGTAAAGTCAAGTCGCTCAACTTGCACGAAA 335  
QY 288 GCCTTGGAGGATCATCCCATCGTCGATGTGAGCTTGATCCTTCGCTATTGTCATC 347  
DB 336 GATTGGTTGGCACCATCCCGTGTGGATTGGTGAAGCTTGACCACTTGTACTTCTGTT 395  
QY 348 TCTGGGTAAATTCATTGTTGGGAGGTACCAAAAGTTTCAGATACGGCTCAAGAGCC 407  
DB 396 TCTGGACAAATTCATTGTTGGTAGGTACCAATAGTTTGCAGATAGATCAAGAGCC 455  
QY 408 TCACACTGACAGCCAGTCACCTCGTATGGTTCATTAAACATGCTATTGTCATGTA--- 464  
DB 456 TCGCCACCGCTGCTGTTCACTAGTATGGCTTTGCTTAACATGCTCATGTCATGTA 515  
QY 465 GCAGTAGAAGACGCTCGATGAGAACCAATACATATCAGGACCAACATAGTGTG 524  
DB 516 GGAACCGAAGAACCTCGACGAAACAAACAAATACAATACATGCTGTAACCAACTGT 575  
QY 525 GATCAGGAGCAACAATGTTGTTCCGGGAATGACACACGCTGCTATCTGGGAATACA 584  
DB 576 GATCTGGGAACGACATGCTGTTCTGGGAAGCAACACAGTATGCTGGGAACACA 635  
QY 585 ACCATGCTCTGGGAGCAACACTGTTGTAACCTGGAAGTGAACAATCTGATGTGTA 644  
DB 636 ACACAGTCTCTGGGAGCAACACCACTTTCATCTGGCAGTGCAATATGTAACGTGA 695  
QY 645 GCACCATGCTGATCAGGACCAAGCATATGTTACTGATAACATATGTTGATCCG 704  
DB 696 GCAACCATATGTTATGTTGGGACCAACAAATATCAATCACTGTAACCAATGACGTATCCG 755  
QY 705 GGAACGCAACAATATGTTCTGGAAGCTTCCATCTGATCAGGGGAGCACAATACCGTAT 764  
DB 756 GCATGATGAAACATGATCTGGAGCTTCCATCTGATCTCGGGAGCCACATCTGTAT 815  
QY 765 CCGGAGCAACAATCTGATCTCGGAGCAACCAATATGATCTGTTGGGAGCAACAAGTCG 824  
DB 816 CTGGAGGTAAACAACACTGATCTGGAAGCAACCACTGCTGATCTGGAAGCAACAAGTCG 875  
QY 825 TAACAGATGTTTATATCTCTAGG-----TGCAAGGATTCCTTCCA 865  
DB 876 TGACAGGAGATGAATGATTTGTGAGGGGATGTTTCCATCTTCCCTTAAAGGAGCTCTAC 935

QY 866 TCTTCCCAAGTTCACTGATGCTTACAATCAATAGATGGAGACAATCAGTTTATGTA 925  
DB 936 CTGTGTCAGTTCACTGATGCTTACAATCACTCTGTTAGGGCAATCGAGTTATGTA 995  
QY 926 CA-CGATATGCACTATCTTTCTCTTTTAAATAA 956  
DB 996 CATGGATATGATCATCACTTTCTCTATAGTTTAA 1027

RESULT 8  
ADM41466  
ID ADM41466 standard; cDNA; 1246 BP.

XX AC ADM41466;  
XX 03-JUN-2004 (first entry)  
XX Tall fescue antifreeze protein cDNA.  
DE Antifreeze; fescue; cold tolerance; transgenic; plant; antitox;  
KW litholytic; nephrotropic; cytotatic; gene; ss.  
XX Schedonorus arundinaceus.

Key Location/Qualifiers  
CDS 53..910  
FT /\*tag= b  
FT /product= "Antifreeze protein"  
FT sig\_peptide 53..103  
FT /\*tag= a  
FT mat\_peptide 104..907  
FT /\*tag= c

WO2004022700-A2.

18-MAR-2004.

09-SEP-2003; 2003WO-NZ000199.

09-SEP-2002; 2002US-0409557P.

(GENE-) GENESIS RES & DEV CORP LTD.  
(WRIG-) WRIGHTSON SEEDS LTD.

Demmer J, Shenk MA, Hall C, Fish SA;

WPI; 2004-248453/23.

P-PSDB; ADM41479.

New antifreeze proteins and encoding polynucleotides, useful for  
modulating cold tolerance in organisms, as food additives, or for  
treating tumors or disorders associated with the presence of unwanted  
biocrystals (e.g. gout).

Claim 1; SEQ ID NO 9; 71pp; English.

The present sequence is that of cDNA encoding an antifreeze protein of  
tall fescue. The cDNA was isolated from a leaf blade cDNA expression  
library. The invention provides forage grass (perennial ryegrass and tall  
fescue) antifreeze proteins and the polynucleotides encoding them  
ADM41458-ADM41483. The polynucleotides were isolated from tissues taken  
at different times of year (winter and spring) and from different parts  
of the plants. The polynucleotides can be used to modulate the cold  
tolerance of an organism, especially plants, mammals, insects, fungi,  
archaea and bacteria. The method involves incorporating an antifreeze  
polynucleotide, under the control of a gene promoter sequence, into the  
genome of the organism, or introducing double-stranded RNA corresponding  
to the polynucleotide into the cells of the organism, thereby inhibiting  
expression of an antifreeze polypeptide. The antifreeze protein can be  
used for the cryopreservation of a cell or tissue, as a food additive of  
a frozen food product, in a method for decreasing the time required to  
dehydrate a composition, to treat a disorder characterised by biocrystals  
associated with disorders such as gout and kidney stones, to preserve the



|                           |     |  |      |
|---------------------------|-----|--|------|
| SQ                        |     | Sequence 1083 BP; 290 A; 257 C; 262 G; 274 T; 0 U; 0 Other;    |      |
| Query Match               |     | 42.8%; Score 419; DB 12; Length 1083;                          |      |
| Best Local Similarity     |     | 70.4%; Pred. No. 1.9e-120;                                     |      |
| Matches 702; Conservative |     | 0; Mismatches 205; Indels 90; Gaps 7;                          |      |
| QY                        | 33  | CCATCGCTGATACATGCGCAAGTGTTCATCGCTGCTGCTCTCTCTGCGGTTCATCTTGC    | 92   |
| Db                        | 45  | CCAGAACCTTAATCCATGCGCAATGTTGGCTGCTGCTCTCTCTCTGCTCTCTCTTGC      | 104  |
| QY                        | 93  | AGGTGCGCAGGAGCAACGTCGTGCTGCTGCCACGACGACCTCCACGCTTGAGGGGCC      | 152  |
| Db                        | 105 | TGGCCATGACGGGAC-----GTGTCGCCACTGGATGACCTCGCGCGCTGCGGGCT        | 158  |
| QY                        | 153 | TGCTGAGAACCTTAAGCGGCAAGAGCGCTCGCGCTCGCGCGCGCATGCTCGCGCGCT      | 212  |
| Db                        | 159 | TGTCGGGAACCTCAATGSC---GGGGGTGCCCTTCTCCGTGGAACATGCTGCTGCTCT     | 215  |
| QY                        | 213 | CATGCTGACGCTGGGAAGTGTGGGATCGGAAACAGCAAGCGCGCGCTGCTGCGGTGC      | 272  |
| Db                        | 216 | CATGCTGCGATTGGGAAGTGTGGGCTCGGATGGTACAAAGCGCGCGCTCACGGCGTTGC    | 275  |
| QY                        | 273 | GGCTCC-----CCAAAGCGCG  | 287  |
| Db                        | 276 | GGCTTCCGATTAGCTCGAGGACTCGGCTAAGCTCAAGTCGCTCAACCTTGCACAGAA      | 335  |
| QY                        | 288 | GCCTTGGAGGAGTATCCCATGCTGCGATTGGTAGCTTGATCACCCTTCGCTATTGGATC    | 347  |
| Db                        | 336 | GATTGGTTGGCACCATCCCGCTGCTGGATTGGTAGCTTGACCAACCATTCGCTACTTGGTTC | 395  |
| QY                        | 348 | TCTCGGGTAACTCATTTGTTGGGAGGTACC-----AAAAGTTTGCAGATAC            | 395  |
| Db                        | 396 | TCTCGGATTAATCATTTGTTGGTAGGACCAATAGTTTGCACATATAGTTTGCAGATA      | 455  |
| QY                        | 396 | GGCTCAAGAGCTCACCACTGACAGCCAGTCACTCGGTATGGTTTCCATTAACTGCTAT     | 455  |
| Db                        | 456 | GACTCAAGGGCTCGCCACCGCTGCTCTCACTAGGTATGGCTTTCGCTTAACATGCCAT     | 515  |
| QY                        | 456 | TGCATGTCA---GCAGTAGAAGACGCTCGATGAGAACCAATACATATTCAGGGACCA      | 512  |
| Db                        | 516 | TGCATGTGAAGGGGAACCGAAGAACCCCTCGACGAAACAAACAAATACATATCATGGACCA  | 575  |
| QY                        | 513 | ACAATAGTTGTTGATCAGGAGCAACAATGTTTTCGGGAATGACAAACGCTGCTGAT       | 572  |
| Db                        | 576 | ACACACTGTTAGATCTGGGACGCAATGTTTCTGGGAACGACAACTGTCTAT            | 635  |
| QY                        | 573 | CTGGGAATPACAAACCATGTGCTGGAGCAACAACACTGTTGTAACTGGAAAGTGACAATA   | 632  |
| Db                        | 636 | GTGGGAACAACAACACTGTGCTGGGAGCAACAACACCATTTGCATCTGGCAGTGACAATA   | 695  |
| QY                        | 633 | CTGTAGTTGGTAGCAACCATGTGCTATCAGGCAAGACATATTTGTTACTGATACATA      | 692  |
| Db                        | 696 | TCGTAATCTGGGAGCAACCATATTTGATGTGGGACCAACATATCATATCTGATACACA     | 755  |
| QY                        | 693 | ATGTTGTATCTCGGGAACCAACAATATGTTCTGGAAGCTTCCATCTGTATCAGGGAGC     | 752  |
| Db                        | 756 | ATGAGCTATCTCGGCAATATATATATGATCTGGAGCTTCCATCTGTATCCGGAGCC       | 815  |
| QY                        | 753 | ACAATACCGTATCCGGGAGCAACAATATCTGTATCCGGGAGCAACAATATCTGTGGA      | 812  |
| Db                        | 816 | ACAATACTGTATCTGGAAGTAAACAACACTGTATCTGGAAGCAACCATGTCTGTATCTGAA  | 875  |
| QY                        | 813 | GCACAAAGTCTGTAACAGATGTTAATATTTCTGTAGG-----TGC                  | 852  |
| Db                        | 876 | GCAACAAAGTCTGTACAGAGATGAATGATTTGTCAGGGGATGCTTCCATCTTCTTAA      | 935  |
| QY                        | 853 | AGGATGCTTCCATCTTCCCAAGTTTCAGTGTAGCTTACAATCAATAGATGAGACATCA     | 912  |
| Db                        | 936 | AGGAGCTCTACCCCTAGTCCAAAGTTCGGTTCGAGCTCACATCACTTGTGAGGACATCG    | 995  |
| QY                        | 913 | CGTTATGTAACCTCA-GGATATGGCATACTTTTCCTT                          | 948  |
| Db                        | 996 | AGTTATGTAACCTTTCATGATATAGCATCATTTCTCCT                         | 1032 |

RESULT 10  
ADM41462  
ID ADM41462 standard; cDNA; 1084 BP.

XX ADM41462;

AC ADM41462;

DT 03-JUN-2004 (first entry)

XX Perennial ryegrass antifreeze protein AFP3 cDNA.

XX Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant;  
antigout; litholytic; nephrotropic; cytostatic; gene; ss.

XX Lolium perenne.

Key Location/Qualifiers  
CDS 60..905  
/\*tag= b  
/product= "Antifreeze protein AFP3"  
sig\_peptide 60..117  
/\*tag= a  
mat\_peptide 118..902  
/\*tag= c

WO2004022700-A2.

18-MAR-2004.

09-SEP-2003; 2003WO-NZ000199.

09-SEP-2002; 2002US-0409557P.

(GENE-) GENESIS RES & DEV CORP LTD.  
(WRIG-) WRIGHTSON SEEDS LTD.

Demmer J, Shenk MA, Hall C, Fish SA;

WPI; 2004-248453/23.

P-PSDB; ADM41475.

New antifreeze proteins and encoding polynucleotides, useful for  
modulating cold tolerance in organisms, as food additives, or for  
treating tumors or disorders associated with the presence of unwanted  
biocrystals (e.g. gout).

Claim 1; SEQ ID NO 5; 71pp; English.

The present sequence is that of cDNA encoding AFP3, an antifreeze protein  
of perennial ryegrass. The cDNA was isolated from a root cDNA expression  
library. The invention provides forage grass (perennial ryegrass and tall  
fescue) antifreeze proteins and the polynucleotides encoding them  
ADM41458-ADM41483. The polynucleotides were isolated from tissues taken  
at different times of year (winter and spring) and from different parts  
of the plants. The polynucleotides can be used to modulate the cold  
tolerance of an organism, especially plants, mammals, insects, fungi,  
archaea and bacteria. The method involves incorporating an antifreeze  
polynucleotide, under the control of a gene promoter sequence, into the  
genome of the organism, or introducing double-stranded RNA corresponding  
to the polynucleotide into the cells of the organism, thereby inhibiting  
expression of an antifreeze polypeptide. The antifreeze protein can be  
used for the cryopreservation of a cell or tissue, as a food additive of  
a frozen food product, in a method for decreasing the time required to  
dehydrate a composition, to treat a disorder characterised by biocrystals  
associated with disorders such as gout and kidney stones, to preserve the  
viability of a molecular biology reagent, to destroy unwanted tissue in a  
patient e.g. tumour tissue, and to protect a plant from damage due to  
frost or freezing.

SQ Sequence 1084 BP; 290 A; 258 C; 262 G; 274 T; 0 U; 0 Other;

Query Match 42.6%; Score 417.4; DB 12; Length 1084;





confectionery. Anti-freeze proteins are especially used in food products, which are heated, e.g. by pasteurisation, blanching or sterilisation prior to freezing. Plants transformed with a nucleic acid sequence encoding the anti-freeze protein have an increased frost tolerance. Prior art anti-freeze proteins have not been applied to commercially available food products, due to high costs and complicated processes for obtaining the protein. Also prior art anti-freeze proteins have tended to destabilise during processing especially during the pasteurisation step. This is overcome by the present anti-freeze protein. The anti-freeze proteins provide an ice particle size following an ice recrystallisation inhibition assay of 15  $\mu$ m or less. The anti-freeze protein ingredient means that mixes can be frozen under quiescent conditions, e.g. in a shop or home freezer without the formation of unacceptable ice crystal shapes and hence with a texture different to products normally obtained via quiescent freezing

Sequence 357 BP; 116 A; 80 C; 89 G; 72 T; 0 U; 0 Other;

|    |     |   |     |
|----|-----|---|-----|
| Qy | 662 | GGGACAAAGCATATCTGTTACTGATACATAATGTTGTATCCGGGAGCGACAAATATGTC   | 721 |
| Db | 181 | GGGACAAACCATATCGTTTACAGACAAACAATTAACGTATCCGGGACGATAATATGTA    | 240 |
| Qy | 722 | TCGTGAAGCTTCCATACTGTATCAGGGGAGCAACAATACCGTATCCGGGAGCAACAATACT | 781 |
| Db | 241 | TCGGGAGCTTTTCATACCGTATCCGGGGGCGCAATACTGTGTCCGGGAGCAACAATACC   | 300 |
| Qy | 782 | GTATCCGGGAGCAACCATATCGTATCTGGGAGCAACAAGTCGTACAGATGCTTTAA      | 838 |
| Db | 301 | GTATCTCGGAGCAACCACTGTTGTTATCTGGAAGCAACAAGTCGTACAGACGCTTTAA    | 357 |

|           |  |
|-----------|--|
| RESULT 14 |  |
| ADM41458  |  |
| ID        | ADM41458 standard; cDNA; 841 BP.                                   |
| XX        |  |
| XX        | ADM41458;  |
| XX        |  |
| XX        | 03-JUN-2004 (first entry)  |
| DT        |  |
| DT        |  |
| XX        |  |
| XX        | Perennial ryegrass antifreeze protein AFP1 cDNA.                   |
| DE        |  |
| XX        |  |
| XX        | Antifreeze; perennial ryegrass; cold tolerance; transgenic; plant; |
| KW        | antigout; litholytic; nephrotropic; cytostatic; gene; ss.          |
| KW        |  |
| XX        |  |
| XX        | Lolium perenne.  |
| XX        |  |
| XX        |  |
| XX        | Location/Qualifiers  |
| PH        | 55..786  |
| FT        | key  |
| FT        | CDS  |
| FT        |  |
| FT        | /*tag= b   |
| FT        | /product= "Antifreeze protein AFP1"                                |
| FT        | 55..114  |
| FT        | /*tag= a   |
| FT        |  |
| FT        | sig_peptide  |
| FT        |  |
| FT        |  |

```

FI      mat_negative      /tag= c
FT      XX
XX      XX
PN      WO2004022700-A2.
XX      XX

```

PD 18-MAR-2004.  
XX  
XX  
PF 09-SEP-2003; 2003WO-NZ000199.  
XX  
PR 09-SEP-2002; 2002US-0409557P.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX (WRIG-) WRIGHTSON SEEDS LTD.  
XX  
XX Demmer J, Shenk MA, Hall C, Fish SA;  
XX  
XX WPI; 2004-248453/23.  
DR P-PSDB; ADM41471.  
DR  
XX  
XX  
PT New antifreeze proteins and encoding polynucleotides, useful for  
PT modulating cold tolerance in organisms, as food additives, or for  
PT treating tumors or disorders associated with the presence of unwanted  
PT biocrystals (e.g. gout).  
XX  
XX Claim 1; SEQ ID NO 1; 71pp; English.  
XX  
XX The present sequence is that of cDNA encoding APPI, an antifreeze protein  
CC of perennial ryegrass. The cDNA was isolated from a pseudostem cDNA  
CC expression library. The invention provides forage grass (perennial  
CC ryegrass and tall fescue) antifreeze proteins and the polynucleotides  
CC encoding them ADM41458-ADM41483. The polynucleotides were isolated from  
CC tissues taken at different times of year (winter and spring) and from  
CC different parts of the plants. The polynucleotides can be used to  
CC modulate the cold tolerance of an organism, especially plants, mammals,  
CC insects, fungi, archaea and bacteria. The method involves incorporating  
CC an antifreeze polynucleotide, under the control of a gene promoter  
CC sequence, into the genome of the organism, or introducing double-stranded  
CC RNA corresponding to the polynucleotide into the cells of the organism,  
CC thereby inhibiting expression of an antifreeze polypeptide. The  
CC antifreeze protein can be used for the cryopreservation of a cell or  
CC tissue, as a food additive of a frozen food product, in a method for  
CC decreasing the time required to dehydrate a composition, to treat a  
CC disorder characterised by biocrystals associated with disorders such as  
CC gout and kidney stones, to preserve the viability of a molecular biology  
CC reagent, to destroy unwanted tissue in a patient e.g. tumour tissue, and  
CC to protect a plant from damage due to frost or freezing.  
XX  
XX Sequence 841 BP; 216 A; 217 C; 222 G; 186 T; 0 U; 0 Other;  
SQ  
Query Match 26.8%; Score 262.2; DB 12; Length 841;  
Best Local Similarity 73.5%; Pred. No. 3.1e-71;  
Matches 349; Conservative 0; Mismatches 123; Indels 3; Gaps 1;  
QY 380 AAAAGTTTGCAGATACGGCTCAAGAGCCTCACCAGTACAGCAGTCACTCGTATGGGT 439  
DB 262 AAGAGCTTCCAGATATTGCTCAAGGGCTCACCGCGCTGCGCTTCACTGGGTAGCG 321  
QY 440 TCCATTACATGCTATTGTCATGTGAGCAGTAG---AAGAACGCTCGATGAAGAACCAAT 496  
DB 322 TTCACTACATGCTCCATTACATGTGAAGCTAGCCAGGAACACTCGAGGAAGACCAAT 381  
QY 497 ACAATATCAGGAGCCAAATAGTGTGGATCAGGAGCAGCAATGTTGTTCCGGGAAAT 556  
DB 382 ACAATACTGGGATCAACAAATGCTGAGATCCGGGAGCAACAATGTTTCTGGGAAC 441  
QY 557 GACACACGGTCTGATCTGGGAATAAACCAATGCTGCGGAGCAACACACTGTTGTA 616  
DB 442 GATAACACTGTCATATCCGGNACACAACTGCTGTCGGGAGCCACACCGTGTGA 501  
QY 617 ACTGAAGTGCATATCTGATGTTGGTAGCAACCATGTCGATCAGGAGCAAAACATTT 676  
DB 502 TTTGGGGTGCACATTCATAGTGGAGTGTACCATGTCGATCTGCGGAACCAACCATGTT 561  
QY 677 GTTACTGATACATAATGTTGTATCCGGGAACGACATATGTTCTGGAAGCTTCCAT 736  
DB 562 GTGACTGACAAAGAAATGCCGTATCCGGGAGCACAAATGCTATCTGGGAAGCCAAAT 621  
QY 737 ACTGTATCAGGGGAGCAAAATACCGTATCCGGGAGCAACAATGCTATCTGCGGAGCAAC 796

DB 622 ACCGTATCCGGAGACCAACCATGATCTGTTGGAGCCACAGTACCGTATCCGGNACAC 681  
QY 797 CATATCGTATCTGGGAGCAACAAAGTGTACAGATGGTTAAATTTCTGTAGGTG 851  
DB 682 AATACGGTATCTGGGAGAACAAATCCGTATATCGGAAACAAATATTTGTATCTG 736

RESULT 15  
ADJ40496  
ID ADJ40496 standard; cDNA; 4536 BP.  
XX  
XX AC ADJ40496;  
XX  
XX 06-MAY-2004 (first entry)  
DT Plant cDNA #1496.  
DE  
XX  
XX Plant; gene; ss; transcription; plant genome augmentation; cereal;  
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;  
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;  
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;  
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;  
KW antifungal.  
XX  
XX Eukaryota.  
XX  
XX US2004016025-A1.  
XX  
XX 22-JAN-2004.  
XX  
XX 26-SEP-2002; 2002US-00260238.  
XX  
XX 26-SEP-2001; 2001US-0325277P.  
PR 26-SEP-2001; 2001US-0325448P.  
PR 04-APR-2002; 2002US-0370620P.  
XX  
XX (BUDW/) BUDWORTH P.  
PA (MOUG/) MOUGHAMER T.  
PA (BRIG/) BRIGGS S P.  
PA (COOP/) COOPER B.  
PA (GLAZ/) GLAZEBROOK J.  
PA (GOFF/) GOFF S A.  
PA (KATA/) KATAGIRI F.  
PA (KREP/) KREPS J.  
PA (PROV/) PROVANT N.  
PA (RICK/) RICHE D.  
PA (ZHUT/) ZHU T.  
XX  
XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;  
PI Goff SA, Katagiri F, Kreps J, Provant N, Ricke D, Zhu T;  
XX  
XX WPI; 2004-190374/18.  
XX  
XX New rice promoter, useful for manipulating crop plants to alter or  
PT improve phenotypic characteristics e.g. produce large quantities of oil  
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance  
PT or high nutritional value.  
XX  
XX Claim 25; SEQ ID NO 1496; 230pp; English.  
XX  
XX The invention relates to plant nucleotide sequences that direct seed-,  
CC leaf- and/or stem-, panicle-, root- or pollen-specific or preferential  
CC or constitutive transcription of an operatively linked nucleic acid  
CC segment. The invention also relates to a method for augmenting a plant  
CC genome and a method of identifying a gene, where its expression is  
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive  
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,  
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,  
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they  
CC encode are useful for manipulating crop plants to alter or improve  
CC phenotypic characteristics, to produce large quantities of oil or  
CC proteins, to incur resistance to insecticides, viruses or fungi, and to

CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants  
CC have a high nutritional value with reduced apical dominance or dwarfism,  
CC early flowering or altered metabolic pathways. This sequence represents a  
CC plant nucleic acid of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification but was obtained in  
CC electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX SQ Sequence 4536 BP; 1102 A; 1247 G; 1221 C; 965 T; 0 U; 1 Other;

```
Query Match      8.1%; Score 79.4; DB 12; Length 4536;
Best Local Similarity 63.1%; Pred. No. 2.2e-13;
Matches 157; Conservative 0; Mismatches 86; Indels 6; Gaps 2;

QY 104 GCAACGTCGTGGTCCACACGACGACCTCCACGCGTTGAGGGGCGCTCGCTGAGAAC 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 GCTGCGGCGCGGCATGCCACCCGGAAGACCTCTCGCGCTGCGGGCATTTGCGGGGAAT 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 164 CT---AAGCGGCAAGAGCGGTCCGCTCCGCGCGCGCATGGTCCGCGCGCTCATGCTGC 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 CTCTCTGCGGGTGGGGCGCGCGGACTCCGCGCGCGTGGTCCGCGTGAACGCTGCTGC 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 221 AGCTGGGAAGTGTGGGATCGAAACAGCAAGCGCGCGCTCGTGGCGTTGCGGCTCCCC 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 GCCTGGGACGCGCTCGCTCGGACGCGC---CGCCGAGTCAAGCGCTGCGGCTCCCC 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 281 AAGCGGCGCTTGGAGGGATCATCCATCGTTCGATTGGTGAGCTTGATCACCTTCGCTAT 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 GGGGAGGTCTCGAGGGGGCCCATCCGCGCTCCCTCGCGCGCTCGCGCGCTCCAGGAC 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 341 TTGGATCTC 349
    ||| ||| |||
Db 313 CTCGACCTC 321
    ||| ||| |||
```

Search completed: May 11, 2005, 00:27:35  
Job time : 657 secs

Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2005, 01:34:03 ; Search time 704 Seconds  
(without alignments)  
8511.756 Million cell updates/sec

Title: US-10-657-852A-3

Perfect score: 980

Sequence: 1 gtttcattcccaatcaaggt.....ccctttacataaaaaaaa 980

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 5654200 seqs, 3057283753 residues

Total number of hits satisfying chosen parameters: 11308400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCIRUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description                             |
|------------|-------|-------------|--------|----|---|
| 1          | 79.4  | 8.1         | 3285   | 18 | US-10-437-963-77805 Sequence 77805, A   |
| 2          | 79.4  | 8.1         | 4536   | 17 | US-10-260-238-1496 Sequence 1496, Ap    |
| 3          | 50.8  | 5.2         | 2883   | 18 | US-10-437-963-4582 Sequence 4582, Ap    |
| 4          | 50.6  | 5.2         | 1912   | 17 | US-10-260-238-568 Sequence 568, Ap      |
| 5          | 50.6  | 5.2         | 2037   | 18 | US-10-437-963-47843 Sequence 47843, A   |
| C 6        | 45.4  | 4.6         | 2498   | 13 | US-10-027-632-102064 Sequence 102064, A |
| C 7        | 45.4  | 4.6         | 2498   | 13 | US-10-027-632-102065 Sequence 102065, A |
| C 8        | 45.4  | 4.6         | 2498   | 13 | US-10-027-632-102066 Sequence 102066, A |
| C 9        | 45.4  | 4.6         | 2498   | 17 | US-10-027-632-102064 Sequence 102064, A |
| C 10       | 45.4  | 4.6         | 2498   | 17 | US-10-027-632-102065 Sequence 102065, A |
| C 11       | 45.4  | 4.6         | 2498   | 17 | US-10-027-632-102066 Sequence 102066, A |

|      |      |     |        |    |   |
|------|------|-----|--------|----|---|
| C 12 | 44.8 | 4.6 | 1614   | 18 | US-10-437-963-79739 Sequence 79739, A   |
| 13   | 44.8 | 4.6 | 3453   | 14 | US-10-101-464A-861 Sequence 861, App    |
| 14   | 44.8 | 4.6 | 3453   | 19 | US-10-864-252-861 Sequence 861, App     |
| 15   | 44.6 | 4.6 | 2589   | 18 | US-10-437-963-47358 Sequence 47358, A   |
| 16   | 43.8 | 4.5 | 235070 | 13 | US-10-087-192-1990 Sequence 1990, Ap    |
| 17   | 43.4 | 4.4 | 485    | 18 | US-10-425-115-155868 Sequence 155868, A |
| C 18 | 43.4 | 4.4 | 1504   | 18 | US-10-719-993-27 Sequence 27, Appl      |
| C 19 | 43.4 | 4.4 | 1569   | 18 | US-10-719-993-26 Sequence 26, Appl      |
| C 20 | 43.4 | 4.4 | 37305  | 18 | US-10-719-993-6767 Sequence 6767, Ap    |
| C 21 | 42.8 | 4.4 | 2379   | 18 | US-10-437-963-7515 Sequence 7515, Ap    |
| C 22 | 42.4 | 4.3 | 632    | 17 | US-10-424-599-138064 Sequence 138064, A |
| C 23 | 42.4 | 4.3 | 990    | 18 | US-10-425-115-105088 Sequence 105088, A |
| 24   | 42.4 | 4.3 | 2485   | 17 | US-10-425-114-24982 Sequence 24982, A   |
| C 25 | 42.2 | 4.3 | 2268   | 18 | US-10-437-963-39626 Sequence 39626, A   |
| C 26 | 42.2 | 4.3 | 2423   | 18 | US-10-425-115-3469 Sequence 3469, Ap    |
| 27   | 41.8 | 4.3 | 492    | 18 | US-10-767-701-22961 Sequence 22961, A   |
| 28   | 41.8 | 4.3 | 3276   | 18 | US-10-437-963-30048 Sequence 30048, A   |
| C 29 | 41.4 | 4.2 | 966    | 18 | US-10-363-345A-3863 Sequence 3863, Ap   |
| C 30 | 41.4 | 4.2 | 966    | 18 | US-10-363-345A-3864 Sequence 3864, Ap   |
| C 31 | 41.4 | 4.2 | 966    | 19 | US-10-363-483A-3863 Sequence 3863, Ap   |
| 32   | 41.4 | 4.2 | 966    | 19 | US-10-363-483A-3864 Sequence 3864, Ap   |
| 33   | 41.4 | 4.2 | 408    | 14 | US-10-101-464A-212 Sequence 212, App    |
| 34   | 41   | 4.2 | 408    | 19 | US-10-864-252-212 Sequence 212, App     |
| 35   | 41   | 4.2 | 667    | 18 | US-10-425-115-67222 Sequence 67222, A   |
| 36   | 40.8 | 4.2 | 311    | 18 | US-10-437-963-47196 Sequence 47196, A   |
| C 37 | 40.8 | 4.2 | 3696   | 18 | US-10-437-963-430 Sequence 430, App     |
| C 38 | 40.4 | 4.1 | 333    | 18 | US-10-437-963-98336 Sequence 98336, A   |
| 39   | 40.4 | 4.1 | 396    | 14 | US-10-101-464A-221 Sequence 221, App    |
| 40   | 40.4 | 4.1 | 396    | 19 | US-10-864-252-221 Sequence 221, App     |
| 41   | 40.4 | 4.1 | 3381   | 10 | US-09-952-267-6 Sequence 6, Appli       |
| C 42 | 40.4 | 4.1 | 12614  | 17 | US-10-242-355-1086 Sequence 1086, Ap    |
| 43   | 40.2 | 4.1 | 1048   | 17 | US-10-424-599-95215 Sequence 95215, A   |
| 44   | 39.8 | 4.1 | 3330   | 18 | US-10-437-963-64573 Sequence 64573, A   |
| 45   | 39.6 | 4.0 | 1671   | 18 | US-10-437-963-7514 Sequence 7514, Ap    |

#### ALIGNMENTS

#### RESULT 1

US-10-437-963-77805  
; Sequence 77805, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 77805  
; LENGTH: 3285  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(3285)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_7766C.1  
US-10-437-963-77805

Query Match 8.1%; Score 79.4; DB 18; Length 3285;  
Best Local Similarity 63.1%; Pred. No. 2.2e-14;  
Matches 157; Conservative 0; Mismatches 86; Indels 6; Gaps 2;

QY 104 GCAAGTCGTGTCGTGTCACACGACGACCTCCACGGTGTGAGGGGCTCGCTGAGAAC 163  
DB 76 GCTGGGGGGGGCCATGACACCCGAGACCTTCCTGGCGCTGCGGGGATTTGCGGGGAAT 135  
QY 164 CT---AAGCGGCAAGAGGACCGTCGCGCTCGCGCGCATGTCGCGGCGCTCATGCTGC 220  
DB 136 CTCTCTGCGGGTGGGGGGCGCGGACTCGCGCGCGGTGGTGGGTGACGCGCTGCTGC 195  
QY 221 AGCTGGAGAGTGTGGGATGCGAAACAGAGAGCGCGCGTGTGGCGCTTGGCGCTCCCC 280  
DB 196 GCCTGGAGCGGCTGCGCTGCGACCGCG---CGCCCGAGTCACGCGCGCTGCGCTCCCC 252  
QY 281 AAGCGCGCTTGGAGGATCATCCATCGATTCGATTTGGTGTGATCATCTTGGCTAT 340  
DB 253 GGGCGAGTCTCGAGGGGCCCATCCCGCTCCCTCGCGCGCTCGCGCTCCAGGAC 312  
QY 341 TTGGATCTC 349  
DB 313 CTCGACCTC 321

## RESULT 2

US-10-260-238-1496

; Sequence 1496, Application US/10260238

; Publication No. US20040016025A1

; GENERAL INFORMATION:

; APPLICANT: Budworth, Paul R.

; APPLICANT: Moughamer, Todd G.

; APPLICANT: Briggs, Steven P.

; APPLICANT: Cooper, Bret

; APPLICANT: Glazebrook, Jane

; APPLICANT: Goff, Stephen A.

; APPLICANT: Katagiri, Fumiyaki

; APPLICANT: Kreps, Joel

; APPLICANT: Provart, Nicholas

; APPLICANT: Ricke, Darrell

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION

; FILE REFERENCE: 60111-NP

; CURRENT APPLICATION NUMBER: US/10/260,238

; PRIOR FILING DATE: 2002-09-26

; PRIOR APPLICATION NUMBER: US 60/325,448

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: US 60/325,277

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: US 60/370,620

; PRIOR FILING DATE: 2002-04-04

; NUMBER OF SEQ ID NOS: 6077

; SEQ ID NO 1496

; LENGTH: 4536

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: N region

; LOCATION: (661)..(661)

; OTHER INFORMATION: n = any nucleotide

US-10-260-238-1496

Query Match 8.1%; Score 79.4; DB 17; Length 4536;  
Best Local Similarity 63.1%; Pred. No. 2.7e-14;  
Matches 157; Conservative 0; Mismatches 86; Indels 6; Gaps 2;

QY 104 GCAAGTCGTGTCGTGTCACACGACGACCTCCACGGTGTGAGGGGCTCGCTGAGAAC 163  
DB 76 GCTGGGGGGGGCCATGACACCCGAGACCTTCCTGGCGCTGCGGGGATTTGCGGGGAAT 135  
QY 164 CT---AAGCGCAAGAGGACCGTCGCGCTCGCGCGCATGTCGCGGCGCTCATGCTGC 220  
DB 136 CTCTCTGCGGGTGGGGGGCGCGGACTCGCGCGCGGTGGTGGGTGACGCGCTGCTGC 195  
QY 221 AGCTGGAGAGTGTGGGATGCGAAACAGAGAGCGCGCGTGTGGCGCTTGGCGCTCCCC 280

DB 196 GCCTGGAGCGGTCGCTGCGACGCGCG---CGCCCGAGTCACGGGCTCGCGCTCCCC 252  
QY 281 AAGCGCGCTTGGAGGATCATCCATCGATTCGATTTGGTGTGATCATCTTGGCTAT 340  
DB 253 GGGCGAGTCTCGAGGGGCCCATCCCGCTCCCTCGCGCGCTCGCGCTCCAGGAC 312  
QY 341 TTGGATCTC 349  
DB 313 CTCGACCTC 321

## RESULT 3

US-10-437-963-4582

; Sequence 4582, Application US/10437963

; Publication No. US2004012343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 4582

; LENGTH: 2883

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_11452C.1

US-10-437-963-4582

Query Match 5.2%; Score 50.8; DB 18; Length 2883;  
Best Local Similarity 56.6%; Pred. No. 5.5e-05;  
Matches 94; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 117 CGTGCCACCACGACGACCTCCACGCGTTGAGGGGCTCGCTGAGAACCTAAGCGGCAAG 176  
DB 104 CTGCGTCCCGAGGAGCGGGCGGCTTCGCGCATCAAGCTGGTTACAGCGGACC 163  
QY 177 GAGCGGTCGCGCTCCGCGCGCATGTGTCGCGCGCTCATGTGAGCTGGGAAGGTGG 236  
DB 164 CCGACGCGCGCTCGCGTGTGGGGCGCGCGCGGACTGCTCGGTGGAGCGGCTCG 223  
QY 237 GATGCGAAACAGAGCGGCGCGTGTGGGGTGGCGGCTCCCA 282  
DB 224 TCTGCGACAAAGCCACCGCGCCAGCTACGAGGCTGAGGCTCCCAA 269

## RESULT 4

US-10-260-238-568

; Sequence 568, Application US/10260238

; Publication No. US20040016025A1

; GENERAL INFORMATION:

; APPLICANT: Budworth, Paul R.

; APPLICANT: Moughamer, Todd G.

; APPLICANT: Briggs, Steven P.

; APPLICANT: Cooper, Bret

; APPLICANT: Glazebrook, Jane

; APPLICANT: Goff, Stephen A.

; APPLICANT: Katagiri, Fumiyaki

; APPLICANT: Kreps, Joel

; APPLICANT: Provart, Nicholas

; APPLICANT: Ricke, Darrell

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION

; FILE REFERENCE: 60111-NP

Db 32 TCCTCTCTCTCGCGGCTTCTTTCGCGCGCGCTCTACTCGCGCGCAACGACGAGGAGCGCCGCGCGCGC 91  
 Qy 144 TGAGGGGCGCTCGCTGAGAACCTTAAGCGGGCAAAAGGAGCCGCTCCGCTCCGCGCGCGCATGGT 203  
 Db 92 TGATGGCGCTCAAGGAGAGCGCTGGACCCGCGCGGGCGGGTGCTGGGGCTCGTGGGCGCGCT 151  
 Qy 204 CCGGCGCTCATGCTGTGAGCTGGGAAGTGTTGGGATGCGAAACAGCAGCAGCGGCGCGCGCTCG 263  
 Db 152 CCGCGCAGCGCTCGCGGGGCTCGTTTCGTTCGGCTGACGTGCGACAGCGGCGGGCGCGCTGA 211  
 Qy 264 TGGCGTTGGCGCTCCCAAGCGCGGCTTGGAGGGATCACTCCCATCGTCCGATTGGTGAAGC 323  
 Db 212 CGGCGCATCTCGCTGCGAGCGCGGGCTCTCCGCGACCTCTCCCTCCGCGCATCGCGGGC 271  
 Qy 324 TTGATCACCTTCGCTATTTTGGATCTCTCGGGTAAATTCATTGGTTGGGGAGGTACCAA 380  
 Db 272 TCCGGCGGCTCACCGGGCTGTACTCTCACTACAAAGCGCATCAAGGGGGCTTATACCCA 328

RESULT 6  
 US-10-027-632-102064/c  
 ; Sequence 102064, Application US/10027632  
 ; Publication No. US20020198371A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 ; POLYMERIZATION: Polymorphisms in the Human Genome  
 ; FILE REFERENCE: 108827.129  
 ; CURRENT APPLICATION NUMBER: US/10/027,632  
 ; CURRENT FILING DATE: 2002-04-30  
 ; PRIOR APPLICATION NUMBER: US 60/218,006  
 ; PRIOR FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 60/198,676  
 ; PRIOR FILING DATE: 2000-04-20  
 ; PRIOR APPLICATION NUMBER: US 60/193,483  
 ; PRIOR FILING DATE: 2000-03-29  
 ; PRIOR APPLICATION NUMBER: US 60/185,218  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/167,363  
 ; PRIOR FILING DATE: 1999-11-23  
 ; PRIOR APPLICATION NUMBER: US 60/156,358  
 ; PRIOR FILING DATE: 1999-09-28  
 ; PRIOR APPLICATION NUMBER: US 60/146,002  
 ; PRIOR FILING DATE: 1999-08-09  
 ; NUMBER OF SEQ ID NOS: 325720  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 102064  
 ; LENGTH: 2498  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-10-027-632-102064

Query Match 4.8%; Score 45.4; DB 13; Length 2498;  
 Best Local Similarity 51.9%; Pred. No. 0.003;  
 Matches 123; Conservative 1; Mismatches 112; Indels 1; Gaps 1;

Qy 600 GCAACACACTGTTGTAACTGGAAGTGACAACTACTGTAGTTGTAGCAACCACTGTCTAT 659  
 Db 346 GAAGCATCCAGAGTGGAAACAGGAAGCATGCGAGTGGAAACAGGAATGTCCAGCATGAA 287  
 Qy 660 CAGGGA-CAAAGCATATTGTTACTGTATAACAATAATGTTGTATCCGCGGAACGACAATAAT 718  
 Db 286 CAGGAAGCAATCCAGATGGAAACAGGAGCATCCAGAGTGGAAACAGGAATGTCCAGCAT 227  
 Qy 719 GTGCTGGAAGCTTCATCTGATATCAGGGAGGACAAATACCGTATCCGGGAGGAAACAAT 778  
 Db 226 GGAACAGGAAGCATCCAGCATGGAACAGGAAGCGTCAAGCGTGGAAACAGGAAGCGTCAAG 167  
 Qy 779 ACTGTATCCGGGAGCAACCATATCGTATCTGGGAGCAACAAGTCTGTAAACAGATGGT 835  
 Db 166 CGTGGAAACAGGAAGCGTCMAGCGTGGAAACAGGAAGCGTCCAGCGTGGAAACAGGAAGT 110

| Query Match           | 4.6%;           | Score 45.4;  | DB 13;    | Length 2498; |
|-----------------------|-----------------|--|-----------|--------------|
| Best Local Similarity | 51.9%;          | Pred. No. 0.003;   |           |              |
| Matches 123;          | Conservative 1; | Mismatches 112;  | Indels 1; | Gaps 1;      |
| Qy                    | 600             | GCAACAACTGTTGTAACCTGGAAGTGACAAATACTGTAGTTGGTAGCAACCATGTCTGAT | 659       |              |
| Db                    | 346             | GAAGCATCCAGATGGNACAGGAAGCATGCAGAGTGGNACAGGAATGTCCAGCATGGAA   | 287       |              |
| Qy                    | 660             | CAGGGA-CAAAGCATATGTTTACTGTATAACAATAATGTTATCCGGGAACGACAATAAT  | 718       |              |
| Db                    | 286             | CAGGAAGCAATCCAGATGGAAACAGGAGCATCCAGAGTGGAAACAGGAATGTCCAGCAT  | 227       |              |
| Qy                    | 719             | GTGTCGTGAAGCTTCCATACTGTATCAGGGGAGCACAATACCGTATCCGGGAGCAACAAT | 778       |              |
| Db                    | 226             | GGAAACAGGAAGCATCCAGCATGGAAACGGNAGCGTCAAGCTGGACAGNAGCGTCAAG   | 167       |              |
| Qy                    | 779             | ACTGTATCCGGGAGCAACCATATCGTATCTGGGAGCAACAAGTCGTAAACAGATGGT    | 835       |              |
| Db                    | 166             | CGTGGAAACAGGAAGCGTCAGCCGTGGAAACAGGAAGCGTCCAGCGTGGAAACAGGAAGT | 110       |              |

```

RESULT 9
US-10-027-632-102064/c
/ Sequence 102064, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 102064
/ LENGTH: 2498
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-102064

```

```

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 102064
; LENGTH: 2498
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-102064

Query Match          4.6%; Score 45.4; DB 17; Length 2498;
Best Local Similarity 51.9%; Pred. No. 0.003;
Matches 123; Conservative 1; Mismatches 112; Indels 1; Gaps 1;

QY      600  GCACCAACTCTGTGTAACTGGGAAGTGACAACTACTGTAGTGTGTAGCAACCATGTCGTAT 659

```

Db 346 GAAGCATCCAGATGGGAACAGGAAGCATGCAGATGGAAACAGGAATGTCCAGCATGGAA 287  
Qy 660 CAGGGA-CAAAGCATATTTACTGATAACAATAATTTCTATCCGGGAACGACAATAAT 718  
Db 286 CAGGAAGCATCCAGATGGAAACAGGAGCATCCAGATGGAAACAGGAATGTCCAGCAT 227  
Qy 719 GTGTCGGAAGCTTCCATCTGTATCAGGGGAGCACAATACCTATCCGGGAGCAACAAT 778  
Db 226 GGAACAGGAAGCATCCAGCATGGAAACAGGAAGCGTCAAGCGTGGAAACAGGAAGCGTCAAG 167  
Qy 779 ACTGTATCCGGGAGCAACCATATCTGTATCTGGAGCAACAATCTGTACAGATGTT 835  
Db 166 CGTGGAAACAGGAAGCGTCMAGCGTGGAAACAGGAAGCGTCCAGCGTGGAAACAGGAAGT 110

## RESULT 10

US-10-027-632-102065/c

; Sequence 102065, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; PRIOR FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 102065

; LENGTH: 2498

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-102065

Query Match 4.6%; Score 45.4; DB 17; Length 2498;  
Best Local Similarity 51.9%; Pred. No. 0.003;  
Matches 123; Conservative 1; Mismatches 112; Indels 1; Gaps 1;

Qy 600 GCAACAACACTGTGTAACTGGGAAGTGACAATACTGTAGTTGTAGCAACCATGTGCTAT 659  
Db 346 GAAGCATCCAGATGGGAACAGGAAGCATGCAGATGGAAACAGGAATGTCCAGCATGGAA 287  
Qy 660 CAGGGA-CAAAGCATATTTACTGATAACAATAATTTCTATCCGGGAACGACAATAAT 718  
Db 286 CAGGAAGCATCCAGATGGAAACAGGAGCATCCAGATGGAAACAGGAATGTCCAGCAT 227  
Qy 719 GTGTCGGAAGCTTCCATCTGTATCAGGGGAGCACAATACCTATCCGGGAGCAACAAT 778  
Db 226 GGAACAGGAAGCATCCAGCATGGAAACAGGAAGCGTCAAGCGTGGAAACAGGAAGCGTCAAG 167  
Qy 779 ACTGTATCCGGGAGCAACCATATCTGTATCTGGAGCAACAATCTGTACAGATGTT 835  
Db 166 CGTGGAAACAGGAAGCGTCMAGCGTGGAAACAGGAAGCGTCCAGCGTGGAAACAGGAAGT 110

## RESULT 11

US-10-027-632-102066/c

; Sequence 102066, Application US/10027632

Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 102066  
LENGTH: 2498  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-102066

Query Match 4.6%; Score 45.4; DB 17; Length 2498;  
Best Local Similarity 51.9%; Pred. No. 0.003;  
Matches 123; Conservative 1; Mismatches 112; Indels 1; Gaps 1;

Qy 600 GCAACAACACTGTGTAACTGGGAAGTGACAATACTGTAGTTGTAGCAACCATGTGCTAT 659  
Db 346 GAAGCATCCAGATGGGAACAGGAAGCATGCAGATGGAAACAGGAATGTCCAGCATGGAA 287  
Qy 660 CAGGGA-CAAAGCATATTTACTGATAACAATAATTTCTATCCGGGAACGACAATAAT 718  
Db 286 CAGGAAGCATCCAGATGGAAACAGGAGCATCCAGATGGAAACAGGAATGTCCAGCAT 227  
Qy 719 GTGTCGGAAGCTTCCATCTGTATCAGGGGAGCACAATACCTATCCGGGAGCAACAAT 778  
Db 226 GGAACAGGAAGCATCCAGCATGGAAACAGGAAGCGTCAAGCGTGGAAACAGGAAGCGTCAAG 167  
Qy 779 ACTGTATCCGGGAGCAACCATATCTGTATCTGGAGCAACAATCTGTACAGATGTT 835  
Db 166 CGTGGAAACAGGAAGCGTCMAGCGTGGAAACAGGAAGCGTCCAGCGTGGAAACAGGAAGT 110

## RESULT 12

US-10-437-963-79739/c

; Sequence 79739, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazov, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 79739

; LENGTH: 1614

; TYPE: DNA



; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 47358

; LENGTH: 2589

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_50134C.1

US-10-437-963-47358

Query Match 4.6%; Score 44.6; DB 18; Length 2589;

Best Local Similarity 51.0%; Pred.No. 0.0057;

Matches 132; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

```

Qy 123 ACCACGACGACCTCCACGCGTTGAGGGGCTCGCTGAGAACCTAAGCGGCAAGGAGCGG 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 80 ACGACTCCGACATGCTCGCGCTGCTCGAGTTCAAGGATGCCATCGCGGACGATCCGCGCG 139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 183 TCGCCCTCCGCGCGCATGCTCGCGCGCTCATGCTGCAGCTGGGAAGGTGTGGGATGCG 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 GCGTCTGAGCTCCTGGAACAAACACCCCTTCTGCGGTGGAATGGGTGAATGCG 199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 243 AAACAGCAAGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 GCGCGCGGAGCATCGCGTCACGCGCTCGAGCTCGCGCGCCAGAACTTGACCGGTGCGAC 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 300 TCATCCCATGCTGATGATGAGCTTGATCACCCTTCGCTATTGGATCTCTCGGGTAATT 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 TCGCCCGCGCTTCTCTCGGGAATCTGAGTTACCTTCACCTGTGACCTTTCTGGTAATC 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 360 CATTTGTTGGGAGGTACC 378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 GCTTCTCTGGGAGATACC 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: May 11, 2005, 04:32:56

Job time : 773 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2005, 23:56:18 ; Search time 4632 Seconds  
(without alignments)  
10251.751 Million cell updates/sec

Title: US-10-657-852A-3  
Perfect score: 980  
Sequence: 1 gtttcattcccaatcaagt.....cccttacataaaaaaaa 980

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 264.2 | 27.0        | 357    | 6     | BD139942    |
| 2          | 264.2 | 27.0        | 357    | 6     | AX019971    |
| 3          | 264.2 | 27.0        | 357    | 8     | LPE277399   |
| 4          | 79.4  | 8.1         | 3392   | 8     | AK121984    |
| 5          | 79.4  | 8.1         | 149371 | 8     | AP004342    |
| 6          | 76    | 7.8         | 349418 | 3     | CR382398    |
| 7          | 74.8  | 7.6         | 2238   | 8     | AK064359    |
| 8          | 74.8  | 7.6         | 2958   | 8     | AK103166    |
| 9          | 70.2  | 7.2         | 176237 | 9     | CNS08CBF    |
| 10         | 67.2  | 6.9         | 169802 | 9     | CNS01DT2    |
| 11         | 63.8  | 6.5         | 646    | 8     | AF372178    |
| 12         | 63.8  | 6.5         | 347582 | 3     | PFMAL4P1    |
| 13         | 62.8  | 6.4         | 300    | 6     | CQ741677    |
| 14         | 60.8  | 6.2         | 261587 | 2     | AC126139    |
| 15         | 59.6  | 6.1         | 172927 | 2     | AC119701    |
| 16         | 58.2  | 5.9         | 290724 | 2     | AC098103    |
| 17         | 57.8  | 5.9         | 231341 | 5     | BX004816    |
| 18         | 55.8  | 5.7         | 243639 | 2     | AC109943    |
| 19         | 55.2  | 5.6         | 342650 | 1     | AP003582    |

|    |      |     |        |    |           |                     |
|----|------|-----|--------|----|-----------|---------------------|
| 20 | 55   | 5.6 | 349418 | 3  | CR382398  | CR382398 Plasmodiu  |
| 21 | 54.2 | 5.5 | 243174 | 2  | AC107550  | AC107550 Rattus no  |
| 22 | 54.2 | 5.5 | 255300 | 2  | AC128061  | AC128061 Rattus no  |
| 23 | 54   | 5.5 | 2000   | 6  | AX655393  | AX655393 Sequence   |
| 24 | 54   | 5.5 | 167601 | 10 | AC144408  | AC144408 Mus muscu  |
| 25 | 54   | 5.5 | 217131 | 10 | AC126053  | AC126053 Mus muscu  |
| 26 | 53.6 | 5.5 | 56310  | 2  | AC084717  | AC084717 Homo sapi  |
| 27 | 53.6 | 5.5 | 222930 | 2  | AC099698  | AC099698 Mus muscu  |
| 28 | 53.6 | 5.5 | 289405 | 2  | AC090126  | AC090126 Mus muscu  |
| 29 | 53.2 | 5.4 | 110000 | 1  | CR522870  | CR522870 Desulfota  |
| 30 | 53.2 | 5.4 | 158901 | 2  | AC145786  | AC145786 xenopus t  |
| 31 | 52.6 | 5.4 | 41951  | 2  | AC139177  | AC139177 Clona sav  |
| 32 | 52.4 | 5.3 | 10029  | 1  | AE010994  | AE010994 Mechanosa  |
| 33 | 52.2 | 5.3 | 109137 | 2  | AC116969  | AC116969 Rattus no  |
| 34 | 52.2 | 5.3 | 172750 | 2  | AC119495  | AC119495 Rattus no  |
| 35 | 52.2 | 5.3 | 217434 | 2  | AC133239  | AC133239 Rattus no  |
| 36 | 52.2 | 5.3 | 236346 | 2  | AC095765  | AC095765 Rattus no  |
| 37 | 52   | 5.3 | 5940   | 8  | THA133651 | THA133651 Trichoder |
| 38 | 51.8 | 5.3 | 1138   | 6  | EO2599    | EO2599 DNA encodin  |
| 39 | 51.8 | 5.3 | 4215   | 9  | HSSIGMG4  | X12843 Human sigma  |
| 40 | 51.8 | 5.3 | 200267 | 10 | AC103606  | AC103606 Mus muscu  |
| 41 | 51.6 | 5.3 | 193223 | 10 | AL672010  | AL672010 Mouse DNA  |
| 42 | 51.6 | 5.3 | 347582 | 3  | PFMAL4P1  | AL034557 Plasmodiu  |
| 43 | 51   | 5.2 | 190316 | 10 | AC122470  | AC122470 Mus muscu  |
| 44 | 51   | 5.2 | 192387 | 10 | AC151578  | AC151578 Mus muscu  |
| 45 | 50.8 | 5.2 | 97900  | 3  | AC087422  | AC087422 Caenorhab  |

ALIGNMENTS

|            |   |        |     |        |                 |
|------------|---|--------|-----|--------|-----------------|
| RESULT 1   | BD139942  | 357 bp | DNA | linear | PAT 18-SEP-2002 |
| LOCUS      | BD139942  |        |     |        |                 |
| DEFINITION | Frozen foods.   |        |     |        |                 |
| ACCESSION  | BD139942  |        |     |        |                 |
| VERSION    | BD139942.1 GI:23234887  |        |     |        |                 |
| KEYWORDS   | JP 2002504316-A/1.  |        |     |        |                 |
| SOURCE     | Lolium perenne  |        |     |        |                 |
| ORGANISM   | Lolium perenne  |        |     |        |                 |
| REFERENCE  | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Poace; Lolium.   |        |     |        |                 |
| AUTHORS    | 1 (bases 1 to 357)  |        |     |        |                 |
| TITLE      | Frozen foods  |        |     |        |                 |
| JOURNAL    | Patent: JP 2002504316-A 1 12-FEB-2002;  |        |     |        |                 |
| COMMENT    | UNILEVER NV<br>OS Lolium perenne (perennial ryegrass)<br>PN JP 2002504316-A/1<br>PD 12-FEB-2002<br>PF 23-DEC-1998 JP 2000528689<br>PR 22-JAN-1998 GB 9801408.7<br>PI CARL DUDLEY JARMAN, CHRISTOPHER MICHAEL SIDEBOTTOM, SARAH TWIGG,<br>PI DAWN WORRAL<br>PC C12N15/09, A01H5/00, A23G9/02, A23L3/375, C07K14/415, C12N15/00 CC<br>Freezed foods<br>FH Key Location/Qualifiers<br>FT source 1..357<br>FT /organism='Lolium perenne (perennial FT |        |     |        |                 |

FEATURES  
source

Location/Qualifiers  
1..357  
/organism="Lolium perenne"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4522"

ORIGIN

Query Match 27.0%; Score 264.2; DB 6; Length 357;  
Best Local Similarity 83.8%; Pred. No. 2.4e-60;  
Matches 299; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 482 GATGAAGAACCAATACATATATCAGGGACCAACAATAGTGTGGATCAGGAGCAACAAT 541

||||| 1 GATGAACGCCGAATACGATTTCTTGGAGCAACATATGTCAGATCGGGAGCAAAAT 60  
542 GTTGTTCGGGAATGACAAACGGTGTGATCTGGGAATAACCAACCATGTGTCTGGGAGC 601  
61 GTTCTTGTGGGAATGACAAACCGCTCATATCTGGGGACAAACATAGTGTGTCTGGGAGC 120  
602 AACACACTGTGTAACTGGAAGTACAACTACTGTAGTTGGTAGCAACCATGTGTCTATCA 661  
121 AACACACTGTGTAACTGGAAGTACAACTACTGTAGTTGGTAGCAACCATGTGTCTATCA 180  
662 GGGCAAAAGCATATTTGTTACTGATTAACAAATATGTTGTATCCGGGAACGCAATAATGTG 721  
181 GGGCAAAAGCATATTTGTTACTGATTAACAAATATGTTGTATCCGGGAACGCAATAATGTG 240  
722 TCTGGAAGCTTCCATATCTGTATGATGAGGAGCAACATACCGTATCCGGGAGCAACATACT 781  
241 TCCGGGAGCTTTCATACCGTATCCGGGGGACAACTACTGTGTCTGGGAGCAACATACC 300  
782 GTATCCGGGAGCAACCATATGTTCTCTGGAGCAACAAAGTCGTAAACAGATGTTAA 838  
301 GTATCTGGGAGCAACCATGTTGTATCTGGAAGCAACAAAGTCGTACAGACGCTTAA 357

RESULT 2  
AX019971  
LOCUS AX019971 357 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 2 from Patent WO9937782.  
ACCESSION AX019971  
VERSION AX019971.1 GI:10043803  
KEYWORDS  
SOURCE Loliium perenne  
ORGANISM Loliium perenne

REFERENCE  
AUTHORS Twigg, S., Worrall, D., Jarman, C. D. and Sidebottom, C. M.  
TITLE Frozen food product  
JOURNAL Patent: WO 937782-A 2 29-JUL-1999;  
TWIGG SARAH (GB); UNILEVER PLC (GB); WORRALL DAWN (GB); JARMAN CARL  
DUDLEY (GB); SIDEBOTTOM CHRISTOPHER MICHAEL (GB); UNILEVER NV (NL)  
Location/Qualifiers

FEATURES  
source  
1..357  
/organism="Loliium perenne"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4522"

## ORIGIN

Query Match 27.0%; Score 264.2; DB 6; Length 357;  
Best Local Similarity 83.8%; Pred. No. 2.4e-60;  
Matches 299; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
QY 482 GATGAAGAACCAATACAAATATCAGGACCAACATAGTGTGGATCAGGAGCAACAAT 541  
DB 1 GATGAAGACCGAATACGATTTCTGGGAGCAACATATCTGTAGTCTGGGAGCAAAAT 60  
QY 542 GTTGTTCGGGAATGACAAACCGTCTGTATCTGGGATACCAACCATGTGTCTGGGAGC 601  
DB 61 GTTCTTGTGGGAATGACAAACCGTCTATCTGGGAGCAACATAGTGTGTCTGGGAGC 120  
QY 602 AACACACTGTGTAACTGGAAGTACAACTACTGTAGTTGGTAGCAACCATGTGTCTATCA 661  
DB 121 AACACACTGTGTAACTGGAAGTACAACTACTGTAGTTGGTAGCAACCATGTGTCTATCA 180  
QY 662 GGGCAAAAGCATATTTGTTACTGATTAACAAATATGTTGTATCCGGGAACGCAATAATGTG 721  
DB 181 GGGCAAAAGCATATTTGTTACTGATTAACAAATATGTTGTATCCGGGAACGCAATAATGTG 240  
QY 722 TCTGGAAGCTTCCATATCTGTATGATGAGGAGCAACATACCGTATCCGGGAGCAACATACT 781  
DB 241 TCCGGGAGCTTTCATACCGTATCCGGGGGACAACTACTGTGTCTGGGAGCAACATACC 300

QY 782 GTATCCGGGAGCAACCATATGTTCTCTGGAGCAACAAAGTCGTAAACAGATGTTAA 838  
DB 301 GTATCTGGGAGCAACCATGTTGTATCTGGAAGCAACAAAGTCGTACAGACGCTTAA 357  
RESULT 3  
LPE277399  
LOCUS LPE277399 357 bp mRNA linear PLN 29-APR-2000  
DEFINITION Loliium perenne partial mRNA for ice recrystallisation inhibition protein.  
ACCESSION AJ277399  
VERSION AJ277399.1 GI:7573547  
KEYWORDS ice recrystallisation inhibition protein.  
SOURCE Loliium perenne  
ORGANISM Loliium perenne

REFERENCE  
AUTHORS Sidebottom, C. M.  
JOURNAL Thesis (1999) University of York  
REFERENCE 2 (bases 1 to 357)  
AUTHORS Sidebottom, C. M.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-2000) Sidebottom C. M., Plant Science, Unilever Research, Colworth House, Sharnbrook, Bedfordshire, MK44 1LQ, UNITED KINGDOM  
COMMENT cds represents presumed mature peptide generated by cleavage before the first amino acid.

## FEATURES

## source

1..357

/organism="Loliium perenne"

/mol\_type="mRNA"

/db\_xref="taxon:4522"

&lt;1..357

/function="freezing tolerance"

/note="ORF1"

/codon\_start=1

/product="ice recrystallisation inhibition protein"

/protein\_id="CA887814.1"

/db\_xref="GI:7573548"

/db\_xref="UniProt/T/EMBL:Q9M3W4"

Translation="DEPNTIGSNNVTVRSGKNVLGNNDNTVIGDNNNSVGSNNVT  
VSGNDNTVGSNHVSGTNIHVTIDNNNNVSGDNNVSGSFHTVSGGHTVSGSNNVT  
GSNHVSGSNNVTDA"

mat\_peptide

1..354

/product="ice recrystallisation inhibition protein"

/note="ORF1"

## ORIGIN

Query Match 27.0%; Score 264.2; DB 8; Length 357;  
Best Local Similarity 83.8%; Pred. No. 2.4e-60;  
Matches 299; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
QY 482 GATGAAGAACCAATACAAATATCAGGACCAACATAGTGTGGATCAGGAGCAACAAT 541  
DB 1 GATGAAGACCGAATACGATTTCTGGGAGCAACATATCTGTAGTCTGGGAGCAAAAT 60  
QY 542 GTTGTTCGGGAATGACAAACCGTCTGTATCTGGGATACCAACCATGTGTCTGGGAGC 601  
DB 61 GTTCTTGTGGGAATGACAAACCGTCTATCTGGGAGCAACATAGTGTGTCTGGGAGC 120  
QY 602 AACACACTGTGTAACTGGAAGTACAACTACTGTAGTTGGTAGCAACCATGTGTCTATCA 661  
DB 121 AACACACTGTGTAACTGGAAGTACAACTACTGTAGTTGGTAGCAACCATGTGTCTATCA 180  
QY 662 GGGCAAAAGCATATTTGTTACTGATTAACAAATATGTTGTATCCGGGAACGCAATAATGTG 721  
DB 181 GGGCAAAAGCATATTTGTTACTGATTAACAAATATGTTGTATCCGGGAACGCAATAATGTG 240  
QY 722 TCTGGAAGCTTCCATATCTGTATGATGAGGAGCAACATACCGTATCCGGGAGCAACATACT 781  
DB 241 TCCGGGAGCTTTCATACCGTATCCGGGGGACAACTACTGTCTCTGGGAGCAACATACC 300

|            |  |  |        |
|------------|--|--|--------|
| Qy         | 782  | GTATCGGGAGCAACATATCTGATCTGCGAGCAACAAAGTCGTAAACAGATGGTTAA | 838    |
| Db         | 301  | GTATCTGGGAGCAACACCTGTTATCTGGAAGCAACAAAGTCGTGACAGCGCTTAA  | 357    |
| RESULT 4   |  |  |        |
| ACCESSION  | AK121984   |  |        |
| LOCUS      | Oryza sativa (japonica cultivar-group) cDNA clone:J033108O10, full insert sequence.  |  |        |
| DEFINITION | AK121984   |  |        |
| VERSION    | AK121984.1   | GI:37991607  |        |
| KEYWORDS   | FLU_CDNA; CAP trapper.   |  |        |
| SOURCE     | Oryza sativa (japonica cultivar-group)   |  |        |
| ORGANISM   | Oryza sativa (japonica cultivar-group)   |  |        |
| REFERENCE  | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.   |  |        |
| AUTHORS    | The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team.  |  |        |
|            | Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ohtsuki, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kuroaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.   |  |        |
| TITLE      | Collection, mapping, and annotation of over 28,000 cDNA clones from Japonica rice  |  |        |
| JOURNAL    | Science  | 301 (5631), 376-379                                      | (2003) |
| MEDLINE    | 22752273   |  |        |
| PUBMED     | 12869764   |  |        |
| AUTHORS    | Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Imotani, K., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kuroaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Oeato, N., Ota, Y., Otonari, Y., Ryu, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A. |  |        |
| TITLE      | Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice  |  |        |
| JOURNAL    | Unpublished  |  |        |
| REFERENCE  | 3 (bases 1 to 3392)  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  |  |        |
| AUTHORS    | Kikuchi, S.  |  |        |
| TITLE      | Direct Submission  |  |        |
| JOURNAL    | Submitted (31-JAN-2003)  |  |        |
| REFERENCE  | Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of   |  |        |



```
/db_xref="GI:23617051"
/translation="MAAMKVVQSGGRQLLSVYLVAVVILPATSSRRPQCIWRE
DRGAVTHLHFDHVDLTDATAVDVNGTGRAFDVAGRGQVVMDDVLTGEPSP
SSPRVGRTOGTVFSDMNPVALLFCNVVLTAGPVAGSTVTLGRDHIHQPLRELSV
GCTGAFMATGVYLNRTASQPRADAVLELDVVFHTRPEYLOSPPPPHRPPPTTV
VW"
15804..16324
/gene="P0585H11.107"
join(<15804..15851,16016..>16324)
/gene="P0585H11.107"
/notes="start and end point are not identified"
join(15804..15851,16016..16324)
/gene="P0585H11.107"
/notes="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC20740.1"
/db_xref="GI:23617052"
/translation="MSRPVGPPTTRASKIRREGRQKGGGEMARSWRRGGGGG
RRGGGKWRSSRASAAATSPDPAGGEATAVEPPDPAATISFACGNDGGGLP
TGDGHPQPSRSLVAA"
complement(20109..20613)
/gene="P0585H11.108"
complement(join(20109..20350,20514..>20613))
/gene="P0585H11.108"
/notes="start and end point are not identified"
complement(join(20109..20350,20514..20613))
/gene="P0585H11.108"
/notes="predicted by FGENSEH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC20741.1"
/db_xref="GI:23617053"
/translation="MSWGCLLLLLMMIVSASSATMIVISNSSSAAAGRRRM
EYSGGCGSVKSTWNERKLCYKGTGCRABGPDYGCYNRPSPFGRFFHVCYSM
NYCKNLIIPSS"
26261..29293
/gene="P0585H11.109"
<26261..>29293
/gene="P0585H11.109"
/notes="start and end point are not identified"
26261..29293
/gene="P0585H11.109"
/notes="contains EST(s): AU056601(S20723),AU056602(S20723)
contains full-length cDNA(s): AK064359"
/codon_start=1
/product="putative phytosulfokine receptor"
/protein_id="BAC20742.1"
Query Match      8.1%; Score 79.4; DB 8; Length 149371;
Best Local Similarity 63.1%; Pred. No. 8.9e-10;
Matches 157; Conservative 0; Mismatches 86; Indels 6; Gaps 2;

Qy 104 GCAACGTGCGTGTGCGACACACAGACCTCCACGCTGAGGGCGCTCGCTGAGAAC 163
Db 26336 GCTGCGCGCGCGCATGCCACCGAAGACCTCTCGCGCTGCGGCGCATTTGCGGGGAAT 26395
Qy 164 CT---AAGCGGCAAGAGCGCTCGCGCTCGCGCGCATGTGCGGCGCTCATCTGCG 220
Db 26396 CTCTCTCGCGGTGGGCGCGCGCGGACCTCGCGCGCGGTGCGGTGACCGCTGCTGCG 26455
Qy 221 AGCTTGGAGAGTGTGGGATCGGAAACAGCAAGCGCGCGCTCGTGGCGTGTGCGGCTCCCC 280
Db 26456 GCTTGGAGCGCGTGTGCTGCGAGCGCG---CGCCGAGTACGCGGCTGCGCTCCCC 26512
Qy 281 AAGCGCGCGCTTGGAGGGATCATCCATCGTCGATTTGGTGGAGCTTGATCATCTTGGCTAT 340
Db 26513 GGGCGAGGTCTCGAGGGGCGCCATCCCGCCCTCTCGCGCCCTCGCGCGCTCCAGGAC 26572
Qy 341 TTGATCTC 349
Db 26573 CTCGACCTC 26581
```

```
RESULT 6
CR382398/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
CDS
```

```
CR382398 349418 bp DNA linear INV 30-MAR-2004
Plasmodium falciparum chromosome 6, complete sequence; segment 1/5.
CR382398 AL844505
CR382398.1 GI:46362238
HTG.
Plasmodium falciparum 3D7
Plasmodium falciparum 3D7
Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
Hall,N., Pain,A., Berriman,M., Churcher,C., Harris,B., Harris,D.,
Mungall,K., Bowman,S., Atkin,R., Baker,S., Barron,A., Brooks,K.,
Buckee,C.O., Burrows,C., Cherevach,I., Chillingworth,C.,
Chillingworth,T., Christodoulou,Z., Clark,L., Clark,R., Corton,C.,
Cronin,A., Davies,R., Davis,P., Dear,P., Dearden,F., Doggett,J.,
Feltwell,T., Goble,A., Goodhead,I., Gwilliam,R., Hamlin,N.,
Hance,Z., Harper,D., Hauser,H., Hornsby,T., Holtroyd,S.,
Horrocks,P., Humphray,S., Jagels,K., James,K.D., Johnson,D.,
Kerhornou,A., Knights,A., Konfortov,B., Kyes,S., Lark,N.,
Lawson,D., Lennard,N., Line,A., Maddison,M., McLean,J., Mooney,P.,
Moule,S., Murphy,L., Oliver,K., Ormond,D., Price,C., Quail,M.A.,
Rabinowitzsch,S., Rajandream,M.A., Rutter,S., Rutherford,K.M.,
Sanders,M., Simmonds,M., Seeger,K., Sharp,S., Smith,R., Squares,R.,
Squares,S., Stevens,K., Taylor,K., Tivey,A., Unwin,I.,
Whitehead,S., Woodward,J., Sulston,J.E., Craig,A., Newbold,C. and
Barrell,B.G.
Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
Nature 419 (6906), 527-531 (2002)
22255708
12368867
2 (bases 1 to 349418)
Cherevach,I., Davis,P., Goodhead,I., Stevens,K., Mungall,K.,
Berriman,M., Pain,A., Hall,N., Atkin,R., Chillingworth,C.,
Doggett,J., Ormond,D., Sanders,M., Hayes,R., Hall,S., Quail,M. and
Barrell,B.G.
Unpublished
3 (bases 1 to 349418)
Cherevach,I., Davis,P., Goodhead,I., Stevens,K., Mungall,K.,
Berriman,M., Pain,A., Hall,N., Atkin,R., Chillingworth,C.,
Doggett,J., Ormond,D., Sanders,M., Hayes,R., Hall,S., Quail,M. and
Barrell,B.G.
Direct Submission
Submitted (20-SEP-2002) P.falciparum Genome Sequencing Consortium,
The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA, UK
4 (bases 1 to 349418)
Cherevach,I., Davis,P., Goodhead,I., Stevens,K., Mungall,K.,
Berry,A.E., Berriman,M., RA Pain,A., Hall,N., Atkin,R.,
Chillingworth,C., Doggett,J., Ormond,D., Sanders,M., Hayes,R.,
Hall,S., Quail,M. and Barrell,B.G.
Direct Submission
Submitted (26-MAR-2004) P.falciparum Genome Sequencing Consortium,
The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA, UK
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/projects/p_falciparum.
Location/Qualifiers
1..349418
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="6"
complement(653..1432)
/gene="VAR pseudogene"
/locus_tag="PFF0005c"
/pseudo
complement(653..1432)
/gene="VAR pseudogene"
/locus_tag="PFF0005c"
/notes="MAL6P1.317"
```



Qy 759 CCGTATCCGGAGCAACAATCTGTATCCGGAGCAACCATATCGTATCTTGGGAGCAACA 818  
 Db 273194 TGGTGGATGATAACAATATGTTGGATGATAACAACATATGTTGGATTAATCAACAATA 273135  
 Qy 819 AAGCTCGTAACAGATGGTTAA 838  
 Db 273134 CTATCGGTAGCGAAGATGAA 273115

RESULT 7  
 AK064359  
 LOCUS  
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:002-108-B08, full insert sequence.  
 ACCESSION AK064359  
 VERSION AK064359.1 GI:32974377  
 KEYWORDS FLI\_CDNA; oligo-capping.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 1 The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ohta, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Iida, Y., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, K., Ueda, M., Matsuura, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from Japonica rice  
 Science 301 (5631), 376-379 (2003)  
 2752273  
 12869764

TITLE 2 (bases 1 to 2238)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsuura, K., Matsuura, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Nami, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Oeato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.  
 Direct Submission  
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@ias.affrc.go.jp)

Tel:81-29-838-7007, Fax:81-29-838-7007  
 This clone is one of the 28K full-length cDNA clones from japonica rice.  
 URL : http://cdna01.dna.affrc.go.jp/cdna/  
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.  
 FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsuura, K. and Murakami, K.  
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.  
 Location/Qualifiers  
 1. 2238  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="002-108-B08"

ORIGIN  
 Query Match 7.6%; Score 74.8; DB 8; Length 2238;  
 Best Local Similarity 68.4%; Pred. No. 7.1e-09;  
 Matches 119; Conservative 0; Mismatches 52; Indels 3; Gaps 1;  
 QY 291 TTGGAGGGATCATCCCATCGTCGATGGTGGAGGTACCAAAAGTTTCAGATACCGCTCAAGAGCTCA 410  
 Db 172 TGGTTGGCACCATCCCTGAATGATCGTGGTCACTTGAATACCTGACCTACTTGGATCTTT 231  
 QY 351 CGGTTAATTCATTGGTTGGGAGGTACCAAAAGTTTCAGATACCGCTCAAGAGCTCA 410  
 Db 232 CAACAAATTCCTTAGTTCGGCGAGATACCGAAGAGTTT---GACGCACTCAAGAGCTTG 288  
 QY 411 CCACTGACAGCGAGTCACCTCGGTATGGTTCATTAACATGCTATTGCATGCA 464  
 Db 289 TCACCGCCGAGCTTCGCCGGTATGGCTTCACTAACATGCGTTGTATGCA 342

RESULT 8  
 AK03166  
 LOCUS  
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033121B13, full insert sequence.  
 ACCESSION AK03166  
 VERSION AK03166.1 GI:32988375  
 KEYWORDS FLI\_CDNA; CAP trapper.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 1 The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,

Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 2958)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotte, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@ias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hiraoka, T., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kouda, M., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ooka, H., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,

Yasunishi, A. and Hayashizaki, Y.  
Location/Qualifiers  
1. 2958  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033121E13"

# ORIGIN

Query Match 7.6%; Score 74.8; DB 8; Length 2958;  
Best Local Similarity 68.4%; Pred. No. 7.5e-09;  
Matches 119; Conservative 0; Mismatches 52; Indels 3; Gaps 1;

QY 291 TTGAGGAGGATCATCCCATCGTCGATGCGTGGAGCTTGCATCCTCGCTTGGATCTCT 350  
DB 1025 TGGTTGGCACCATCCCTGGAATGATCGTCGCTCACTTGAACCTGACCTTCTT 1084  
QY 351 CGGGTAAATTCATTGTTGGGGAGGTACCAAAAGTTTGCAGATACGGCTCAAGAGCCTCA 410  
DB 1085 CAAACAATCTTGTAGTCGGCGAGATACCGAAGGTTT---GACGAGCTCAAGAGCCTTG 1141  
QY 411 CCACTGACAGCAGCTACTCGTATGGGTTCCATTAACATGCTATTGTCATGTGA 464  
DB 1142 TCACCGCCGACGTTGCGCGGTATGGGTTCACTAATACATGCGCTTGTATGTGA 1195

# RESULT 9

CNS08CBF/c

LOCUS

DEFINITION Human chromosome 14 DNA sequence BAC R-731P5 of library RPCI-11 from chromosome 14 of Homo sapiens (Human), complete sequence.

ACCESSION

AL928742

VERSION

AJ928742.3

KEYWORDS

HTG.

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

1 (bases 1 to 176237)

2 (bases 1 to 176237)

Genoscope.

Direct Submission

Genoscope.

Submitted (19-SEP-2002) Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - FRANCE (E-mail : [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr))

Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

On Sep 24, 2002 this sequence version replaced gi:23306222.

Genome Center

Center: Genoscope / Centre National de Sequencage

Center code: GS

Web site: <http://www.genoscope.cns.fr/>

Contact: [Secref@genoscope.cns.fr](mailto:Secref@genoscope.cns.fr)

-----

The following sequence is oriented from the T7 to the SP6 end.

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

```

10 - 19 : 48
20 - 29 : 165
30 - 39 : 825
40 - 49 : 3919
50 - 59 : 6201
60 - 69 : 8574
70 - 79 : 22106
80 - 89 : 54448
90 - 99 : 79950
-----
Percentage of bases with a quality value >= 40 : 99 %.

FEATURES
source
Location/Qualifiers
1. .176237
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-731P5"
/clone_lib="RPCI-11"
14470..14567
/notes="matching EMBL:G14654
RHdb:RH7974
dbSTS:STS21065
Identified using the e-PCR software (G. Schuler)"
57606..57752
/notes="matching EMBL:L00022
RHdb:RH69981
dbSTS:STS48887
Identified using the e-PCR software (G. Schuler)"
60040..60260
/notes="matching EMBL:M55420
RHdb:RH80278
dbSTS:STS56700
Identified using the e-PCR software (G. Schuler)"
69316..69421
/notes="matching EMBL:H26187
RHdb:RH44850
dbSTS:STS37918
Identified using the e-PCR software (G. Schuler)"
145938..146060
/notes="matching EMBL:Z53040
RHdb:RH9751
dbSTS:STS5550
Identified using the e-PCR software (G. Schuler)"
158529..158653
/notes="matching EMBL:RA160692
RHdb:RH48485
dbSTS:STS41544
Identified using the e-PCR software (G. Schuler)"

ORIGIN
Query Match 7.2%; Score 70.2; DB 9; Length 176237;
Best Local Similarity 56.5%; Pred. No. 2.8e-07;
Matches 130; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 600 GCACCAACACTGTTGTAATCGGAAGTGACAAATCTGTAGTTGGTAGCAACCATGCGTAT 659
DB 7784 GAAGCATCCAGATGGNACAGGAATGTCAGCATGGACAGGAAGCATCCAGAAATGTA 7725
QY 660 CAGGGGACAAAGATATGTTACTGTGATAACAATAATGTTGATCCGGGAACGACAATAATG 719
DB 7724 CAGGAAGCATCCAGATGGGAACAGGAATGTCAGCATGGACAGGAAGCATCAAGATG 7665
QY 720 TGTCTGGAGCTCCATCATGTTATCAGGGGAGCACATACCGTATCCGGGAGCAACAATA 779
DB 7664 TAACAGGAAGCATCCAGATGGGAACAGGAATGTCAGCATGGACAGGAAGCATCCAGA 7605
QY 780 CTGTATCCGGGAGCAACATATCGTATCTCGGAGCAACAAGTCGTAACA 829
DB 7604 ATGTAACAGGAAGCATCCAGATGGACAGGAAGCATCCAGATGGNACA 7555

RESULT 10

```

CNS01DT2 169802 bp DNA linear PRI 19-NOV-2001  
Human chromosome 14 DNA sequence BAC R-417P24 of library RPCI-11  
from chromosome 14 of Homo sapiens (Human), complete sequence.  
AL122127  
AL122127.6 GI:17026193  
HTG.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 169802)  
Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,  
Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,  
Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,  
Gyapay,G., Saurin,W. and Weissbach,J.  
Sequencing of the human chromosome 14  
Unpublished  
2 (bases 1 to 169802)  
Genoscope.  
Direct Submission  
Submitted (19-NOV-2001) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
On Nov 20, 2001 this sequence version replaced gi:14715169.  
-----  
Center: Genoscope / Centre National de Sequencage  
Center code: GS  
Web site: http://www.genoscope.cns.fr/  
Contact: SeqRef@genoscope.cns.fr  
-----  
The following BAC sequence is oriented from the T7 to the SP6 end.  
-----  
Assembly program: Phrap; version 2.0  
Quality coverage: 7.56x in Q20 bases; sum-of-contigs  
-----  
Overall quality chart :  
Range : bases  
0 :  
1 - 9 :  
10 - 19 :  
20 - 29 : 11  
30 - 39 : 159  
40 - 49 : 4325  
50 - 59 : 11964  
60 - 69 : 13264  
70 - 79 : 28113  
80 - 89 : 55713  
90 - 99 : 56253  
-----  
Percentage of bases with a quality value >= 40 : 99 %.

FEATURES  
source  
Location/Qualifiers  
1. .169802  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="14"  
/clone="R-417P24"  
/clone\_lib="RPCI-11"  
11427..11641  
/note="matching EMBL:G33053  
RHdb:RH67749  
dbSTS:STS47676  
Identified using the e-PCR software (G. Schuler)"  
11439..11536  
/note="matching EMBL:G14654  
RHdb:RH7974  
dbSTS:STS21065  
Identified using the e-PCR software (G. Schuler)"  
39191..39405  
/note="matching EMBL:G33053  
RHdb:RH67749







Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartabeyn, A., Sisson, J., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steindler, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villabana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 261587)  
Worley, K.C.

Direct Submission  
Submitted (04-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 261587)

Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24819414.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: GHU

Center clone name: CH230-104014

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 245312 bases at least Q40

Consensus quality: 249138 bases at least Q30

Consensus quality: 252151 bases at least Q20

Estimated insert size: 257796; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
\* NOTE: This sequence may represent more than one clone.  
\* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 250010: contig of 250010 bp in length

\* 250011 250110: gap of unknown length

\* 250111 253467: contig of 3357 bp in length

\* 253468 253567: gap of unknown length

\* 253568 255148: contig of 1581 bp in length  
\* 255149 255248: gap of unknown length  
\* 255249 256807: contig of 1559 bp in length  
\* 256808 256907: gap of unknown length  
\* 256908 258585: contig of 1678 bp in length  
\* 258586 258685: gap of unknown length  
\* 258686 260166: contig of 1481 bp in length  
\* 260167 261587: contig of 1321 bp in length.  
\* 260267 Location/Qualifiers

## FEATURES

## source

1..261587  
/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-104014"

1..2511

/note="wgs end\_extension"

clone\_end:T7"

2562..3793

/note="wgs end\_extension"

clone\_end:T7"

6414..7318

/note="clone\_boundary"

clone\_end:T7"

site:ECORI

end\_sequence:BH325436"

complement(246622..247488)

/note="clone\_boundary"

clone\_end:Sp6

site:ECORI

end\_sequence:BH325437"

## ORIGIN

Query Match 6.2%; Score 60.8; DB 2; Length 261587;

Best Local Similarity 51.9%; Pred. No. 0.00011;

Matches 165; Conservative 0; Mismatches 147; Indels 6; Gaps 1;

QY 512 AACATAGTGTGGATCAGGAGCAACAATGTTGTTCCGGGAATGACACACGGTGTA 571

Db 83753 ATCAGCACTGTTGTAGCATGTATCAGCACTGTTGTAGCATGTATCAGCACTTTTGTAGCA 83694

QY 572 TCTGGGAATACACCACTGTCTGGGAGCAACAACACTGTTGTAACAGTGAAGTGAAT 631

Db 83693 T-----GTATCAGCACTGTGTTAAACATGTATCAGCACTGTTGTAGCATGTATCAGCACT 83640

QY 632 ACTGTAGTTGTAGCAACCATGTCGTATCAGGAGCAACAACATATCTTACTGATAACAAT 691

Db 83639 GTTGTAGCATGTATCAGCACTGTTGTAGCATGTATCAGCACTGTGTTAATATGTATCAGC 83580

QY 692 AATGTTGTATCCGGGAACGACAATAATGTGTCTGGAAGCTTCCATACTGTATCAGGGGAG 751

Db 83579 ACTGTTGTAGCATGTATCAGCACTGTTTAAACATGTATCAGCACTGTTGTAGCATGTATC 83520

QY 752 CACATACCGTATCCGGGAGCAACAATCTGTATCCGGGAGCAACCATATCTGTCGTTGGG 811

Db 83519 AGCACTTTTGTAGCATGTATCAGCACTGTTGTAGCATGTATCAGCACTGTTGTAGCATGT 83460

QY 812 AGCAACAAAGTCGTAACA 829

Db 83459 ATCAGCACTGTTGTATACA 83442

## RESULT 15

AC119701/c

LOCUS

DEFINITION

AC119701

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AC119701 Rattus norvegicus clone CH230-462N7, WORKING DRAFT SEQUENCE.

AC119701.7 GI:25137804

HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;



```
Db 172320 AGCATGTATCAGCACTGGGGTAGCATGTATCAGCACTGTTGTAGCATGTATCAGCACTTT 172261
Qy 760 CGTATCCGGGAGCAACAATACTGTATCCGGGAGCAACCATATCGTATCTGGGAGCAACAA 819
Db 172260 TGTAGCATGTATCAGCACTGTTGTAGCATGTATCAGCACTGTTGTAACATGTATCAGCAT 172201
Qy 820 AGTCGTAACA 829
Db 172200 TGGTATAACA 172191
```

Search completed: May 11, 2005, 02:51:19  
Job time : 4640 secs

*This Page Blank (uspto)*